# PCT ) W

#### WELTORGANISATION FÜR GEISTIGES EIGENTUM Internationales Büro



INTERNATIONALE ANMELDUNG VERÖFFENTLICHT NACH DEM VERTRAG ÜBER DIE INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES PATENTWESENS (PCT)

(51) Internationale Patentklassifikation 7:
C07K 14/00
A2
(11) Internationale Veröffentlichungsnummer: WO 00/17233
(43) Internationales
Veröffentlichungsdatum: 30. März 2000 (30.03.00)

(21) Internationales Aktenzeichen: PCT/EP99/07055 (81) Bestin

(22) Internationales Anmeldedatum: 22. September 1999

(22.09.99)

(30) Prioritätsdaten:

198 43 279.8 22, September 1998 (22.09.98) DE 199 23 567.8 21, Mai 1999 (21.05.99) DE

(71)(72) Anmelder und Erfinder: JOMAA, Hassan [DE/DE]; Breslauer Strasse 24, D-35398 Gießen (DE).

(74) Anwälte: PANTEN, Kirsten usw.; Reichel und Reichel, Parkstrasse 13, D-60322 Frankfurt am Main (DE). (81) Bestimmungsstaaten: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO Patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), eurasisches Patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), europäisches Patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI Patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

#### Veröffentlicht

Ohne internationalen Recherchenbericht und erneut zu veröffentlichen nach Erhalt des Berichts.

(54) Title: GENES OF THE 1-DESOXY-D-XYLULOSE BIOSYNTHETIC PATHWAY

(54) Bezeichnung: GENE DES 1-DESOXY-D-XYLULOSE-BIOSYNTHESEWEGS

#### (57) Abstract

The invention relates to the 1-desoxy- D-xylulose- 5-phosphate reductoisomerase gene, the 1-desoxy- D-xylulose- 5-phosphate-synthase gene and the gcpE gene of the 1-desoxy- D-xylulose biosynthetic pathway and to their use for transforming vectors, host organisms and plants and for determining substances that inhibit this biosynthetic pathway.

#### (57) Zusammenfassung

Die vorliegende Erfindung betrifft das 1-Desoxy- D-xylulose- 5-phosphatreduktoisomerase -Gen, das 1-Desoxy- D-xylulose- 5-phosphat- Synthase- Gen und das gepE-Gen des 1-Desoxy- D-xylulose- Biosynthesewegs und ihre Verwendung zur Transformation von Vektoren, Wirtsorganismen und Pflanzen und zur Bestimmung von Stoffen, die diesen Biosyntheseweg inhibieren.

### --- TEDIGLICH ZUR INFORMATION

Codes zur Identifizierung von PCT-Vertragsstaaten auf den Kopfbögen der Schriften, die internationale Anmeldungen gemäss dem PCT veröffentlichen.

AL	Albanien	ES	Spanien	LS	Lesotho	. SI	Slowenien
AM	Armenien	FT	Finnland	LT	Litauen	SK	Slowakei
AT	Osterreich	FR	Frankreich	LU	Luxemburg	SN	Senegal
AU	Australien	GA	Gabun	LV	Lettland	SZ	Swasiland
AZ	Ascrbaidschan	GB	Vereinigtes Königreich	MC	Monaco	TD	Tschad
BA	Bosnien-Herzegowina	GE	Georgien	MD	Republik Moldau	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagaskar	TJ	Tadschikistan
BE.	Belgien	GN	Guinea	MK	Die ehemalige jugoslawische	TM .	Turkmenistan
	•	GR	Griechenland		Republik Mazedonien	TR	Turkei
BF	Burkina Faso	HU	•	ML	Mali	TT	Trinidad und Tobago
BG	Bulgarien	IE.	Ungara iriand	MN	Mongolei	UA	Ukraine
BJ	Benin		*	MR	Mauretanien	UG	Uganda
BR	Brasilien	IL	Israel	MW	Malayi	US	Vereinigte Staaten von
BY	Belarus	18	Island	MX	Mexiko .		Amerika
CA	Kanada	IT	Italien			UZ	Usbekistan
CF	Zentralafrikanische Republik	JP	Japan	NE	Niger	VN	Vietnam
CG	Kongo	KE	Kenis	NL	Niederlande		Jugoslawico
CH	Schweiz	KG	Kirgisistan	NO	Norwegen	YU	Zimbabwe
CI	Côte d'Ivoire	KP	Demokratische Volksrepublik	NZ	Neusceland	ZW	Zimozowe
CM	Kamerun		Korea	PL	Polen		
CN	China	KR	Republik Korea	PT	Portugal		
CU	Kuba	кz	Kasachstan	RO	Ruminien .		•
CZ	Tschechische Republik	LC	St. Lucia	RU	Russische Föderation		
DE	Deutschland	u	Liechtenstein	SD	Sudan		
DK	Dänemark	LK	Sri Lanka	SE	Schweden		
EE	Estland	LR	Liberia	SG	Singapur		•
_							

## Gene des 1-Desoxy-D-xylulose-Biosynthesewegs

Die vorliegende Erfindung betrifft DNA-Sequenzen, die bei Integration in das Genom von Viren, Eukaryonten und Prokaryonten die Isoprenoid-Biosynthese verändern sowie gentechnologische Verfahren zur Herstellung dieser transgenen Viren, Eukaryonten und Prokaryonten. Außerdem betrifft sie Verfahren zur Identifiziereung von Stoffen mit herbizider, antimikrobieller, antiparasitärer, antiviraler, fungizider, bakterizider Wirkung bei Pflanzen und antimikrobieller, antiparasitärer, antimykotischer, antibakterieller und antiviraler Wirkung bei Mensch und Tier.

Der Biosyntheseweg zur Bildung von Isoprenoiden über den klassischen Acetat/ Mevalonat-Weg und einen alternativen, Mevalonat-unabhängigen Biosyntheseweg, den Desoxy-D-xylulose-Phosphat-Weg, ist bereits bekannt (Rohmer, M., Knani, M., Simonin, P., Sutter, B., and Sahm, H. (1993): Biochem. J. 295: 517-524).

Es ist aber nicht bekannt, wie und über welche Wege in Viren, Eukaryonten und Prokaryonten eine Änderung der Isoprenoidkonzentration über den Desoxy-D-xylulose-Phoshat-Weg erreicht werden kann. In Fig. 1 ist dieser Biosyntheseweg dargestellt.

Es werden daher DNA-Sequenzen zur Verfügung gestellt, die für die 1-Desoxy-D-xylulose-5-phosphat-Synthase (DOXP-Synthase),: 1-Desoxy-D-xylulose-5-phosphatreduktoisomerase(DOXP-Reduktoisomerase) oder das gcpE-Protein kodieren. Alle drei Gene und Enzyme sind an der Isoprenoid-Biosynthese beteiligt.

Das gcpE-Protein hat eine Kinasefunktion und katalysiert die Phosphorylierung eines Zuckers oder eines Phosphorzuckers oder einer Vorstufe der Isoprenoidbiosynthese, insbesondere die Phosphorylierung von 2-C-Methyl-D-erythritol, 2-C-Methyl-D-erythritol-phosphat, insbesondere 2-C-Methyl-D-erythritol-4-phosphat, 2-C-Methyl-D-erythrose, 2-C-Methyl-D-erythrose-

phosphat, insbesondere 2-C-Methyl-D-erythrose-4-phosphat. In der Vorstufe der Isoprenoidsynthese katalysiert das gcpE-Protein insbesondere die Phosphorylierung der folgenden Substanzen:

 $\begin{array}{l} \text{CH}_2\left(\text{OH}\right) - \text{C}\left(\text{CH}_3\right) = \text{C}\left(\text{OH}\right) - \text{CH}_2 - \text{O} - \text{PO}\left(\text{OH}\right)_2, \quad \text{CH}_2\left(\text{OH}\right) - \text{C}\left(\text{CH}_3\right) = \text{C}\left(\text{OH}\right) - \text{CH}_2 - \text{OH}, \\ \text{CH}_2\left(\text{OH}\right) - \text{CH}\left(\text{CH}_3\right) - \text{CO} - \text{CH}_2 - \text{O} - \text{PO}\left(\text{OH}\right)_2, \quad \text{CH}_2 = \text{C}\left(\text{CH}_3\right) - \text{CO} - \text{CH}_2 - \text{OH}, \\ \text{CH}_2 = \text{C}\left(\text{CH}_3\right) - \text{CO} - \text{CH}_2 - \text{O} - \text{PO}\left(\text{OH}\right)_2, \quad \text{CH}_2 = \text{C}\left(\text{CH}_3\right) - \text{CH}\left(\text{OH}\right) - \text{CH}_2 - \text{OH}, \\ \text{CH}_2 = \text{C}\left(\text{CH}_3\right) - \text{CH}\left(\text{OH}\right) - \text{CH}_2 - \text{O} - \text{PO}\left(\text{OH}\right)_2, \quad \text{CH}_2 = \text{C}\left(\text{CH}_3\right) - \text{CH}\left(\text{OH}\right) - \text{CH}_2 - \text{OH}, \\ \text{CH}_2\left(\text{OH}\right) - \text{C}\left(\text{CH}_2\right) - \text{C}\left(\text{OH}\right) - \text{CH}_2 - \text{O} - \text{PO}\left(\text{OH}\right)_2, \quad \text{CHO} - \text{CH}\left(\text{CH}_3\right) - \text{CH}\left(\text{OH}\right) - \text{CH}_2 - \text{OH}, \\ \text{CHO} - \text{CH}\left(\text{CH}_3\right) - \text{CH}\left(\text{OH}\right) - \text{CH}_2 - \text{O} - \text{PO}\left(\text{OH}\right)_2, \quad \text{CH}_2\left(\text{OH}\right) - \text{C}\left(\text{OH}\right) \left(\text{CH}_3\right) - \text{CH}\left(\text{OH}\right) - \text{CH}_2 - \text{OH}, \\ \text{CH}\left(\text{OH}\right) = \text{C}\left(\text{CH}_3\right) - \text{CH}\left(\text{OH}\right) - \text{CH}_2 - \text{O} - \text{PO}\left(\text{OH}\right)_2, \quad \text{CH}\left(\text{OH}\right) = \text{C}\left(\text{CH}_3\right) - \text{CH}\left(\text{OH}\right) - \text{CH}_2 - \text{OH}, \\ \text{CH}_3\right)_2 + \text{C} - \text{CO} - \text{CH}_2 - \text{O} - \text{PO}\left(\text{OH}\right)_2, \quad \text{CH}_3\right)_2 + \text{C} - \text{CO} - \text{CH}_2 - \text{O} - \text{H}. \\ \text{(CH}_3)_2 + \text{C} - \text{C} + \text{C} + \text{OH} - \text{C} + \text{C} - \text{C} - \text{OH} - \text{C} + \text{C} - \text{C} - \text{C} - \text{C} + \text{C} - \text{C} - \text{C} + \text{C} - \text{C} - \text{C} - \text{C} + \text{C} - \text{C} -$ 

Die DOXP-Synthase katalysiert die Kondensation von Pyruvat und Glyceraldehyd-3-phosphat zu 1-Deoxy-D-xylulose-5-phosphat und die DOXP-Reduktoisomerase katalysiert die Umwandlung von 1-Deoxy-D-xylulose-5-phosphat zu 2-C-Methyl-D-erythritol-4-phosphat. (siehe Fig. 1).

Die Erfindung betrifft die folgenden DNA-Sequenzen:
DNA-Sequenzen, die für ein Polypeptid mit der in SEQ ID NO:2
DNA-Sequenzen, die für ein Polypeptid mit der in SEQ ID NO:2
dargestellten Aminosäuresequenz codieren oder für ein Analoges
oder Derivat des Polypeptids gemäß SEQ ID NO:2, worin eine oder
mehrere Aminosäuren deletiert, hinzugefügt oder durch andere
Aminosäuren substituiert worden sind,

DNA-Sequenzen, die für ein Polypeptid mit der in SEQ ID NO:4 dargestellten Aminosäuresequenz codieren oder für ein Analoges oder Derivat des Polypeptids gemäß SEQ ID NO:4, worin eine oder mehrere Aminosäuren deletiert, hinzugefügt oder durch andere Aminosäuren substituiert worden sind,

sowie DNA-Sequenzen, die für ein Polypeptid mit der in SEQ ID NO: 6 dargestellten Aminosäuresequenz codieren oder für ein Analoges oder Derivat des Polypeptids gemäß SEQ ID NO: 6, worin eine oder mehrere Aminosäuren deletiert, hinzugefügt oder durch andere Aminosäuren substituiert worden sind.

Die Gene und ihre Genprodukte (Polypeptide) sind im Sequenzprotokoll mit ihrer Primärstruktur aufgeführt und haben folgende Zuordnung:

SEQ ID NO:1: 1-Desoxy-D-xylulose-5-phosphatreduktoisomerase-Gen

SEQ ID NO:2: 1-Desoxy-D-xylulose-5-phosphatreduktoisomerase

SEQ ID NO:3: 1-Desoxy-D-xylulose-5-phosphat-Synthase-Gen

SEQ ID NO:4: 1-Desoxy-D-xylulose-5-phosphat-Synthase

SEQ ID NO:5: gcpE-Gen

SEQ ID NO:6 : gcpE-Proteine.

Die DNA-Sequenzen stammen alle aus Plasmodium falciparum.

Außer den im Sequenzprotokoll genannten DNA-Sequenzen sind auch solche geeignet, die infolge der Degeneration des genetischen Codes eine andere DNA-Sequenz besitzen, jedoch für das gleiche Polypeptid oder für ein Analoges oder Derivat des Polypeptids kodieren, worin eine oder mehrere Aminosäuren deletiert, hinzugefügt oder durch andere Aminosäuren substituiert worden sind.

Die erfindungsgemäßen Sequenzen eignen sich für die Expression von Genen in Viren, Eukaryonten und Prokaryonten, die für die Isoprenoid-Biosynthese des 1-Desoxy-D-xylulose-Wegs verantwortlich sind.

Erfindungsgemäß gehören zu den Eukaryonten oder eukaryontischen Zellen tierischen Zellen, Pflanzenzellen, Algen, Hefen, Pilze und zu den Prokaryonten oder prokaryontischen Bakterien Archaebakterien und Eubakterien.

Bei Integration einer DNA-Sequenz in ein Genom, auf der eine der oben angegebenen DNA-Sequenzen lokalisiert ist, wird die Expression der oben beschriebenen Gene in Viren, Eukaryonten und Prokaryonten ermöglicht. Die erfindungsgemäß transformierten Viren, Eukaryonten und Prokaryonten werden in an sich bekannter Weise gezüchtet und das währenddessen gebildete Isoprenoid isoliert und gegebenenfalls gereinigt. Nicht alle Isoprenoide müssen isoliert werden, da die Isoprenoide in einigen Fällen direkt in die Raumluft abgegeben werden.

Die Erfindung betrifft ferner ein Verfahren zur Herstellung von transgenen Viren, Eukaryonten und Prokaryonten zur Veränderung des Isoprenoid-Gehaltes, das die folgenden Schritte enthält.

- a) Herstellung einer DNA-Sequenz mit folgenden Teilsequenzen
  - i) Promotor, der in Viren, Eukaryonten und Prokaryonten aktiv ist und die Bildung einer RNA im vorgesehenen Zielgewebe oder den Zielzellen sicherstellt,
  - ii) DNA-Sequenz, die für ein Polypeptid mit der in SEQ ID NO:2,4 oder 6 dargestellten Aminosäuresequenz codieren oder für ein Analoges oder Derivat des Polypeptids gemäß SEQ ID NO:2,4 oder 6,
  - iii) 5'- und 3'-nichttranslatierte Sequenz, die in Viren, Eukaryonten und Prokaryonten die Expression der bezeichneten Gene ermöglichen oder verbessern,
- b) Transfer und Einbau der DNA-Sequenz in das Genom von Viren, prokaryontischen oder eukaryontischen Zellen mit oder ohne Verwendung eines Vektors (z.B. Plasmid, virale DNA).

Aus derart transformierten Pflanzenzellen können die intakten ganzen Pflanzen regeneriert werden.

Die für die Proteine kodierenden Sequenzen mit den Nukleotidabfolgen Seq ID NO:1, Seq ID NO:3 und Seq ID NO: 5 können mit einem die Transkription in bestimmten Organen oder Zellen sicherstellenden Promotor versehen werden, der in sense-Orientierung (3'-Ende des Promotors zum 5'-Ende der kodierenden Sequenz) an die Sequenz, die das zu bildende Protein kodiert, gekoppelt ist. An das 3'-Ende der kodierenden Sequenz wird ein die Termination der mRNA-Synthese bestimmendes Terminationssignal angehängt. Um das zu exprimierende Protein in bestimmte subzelluläre Kompartimente, wie Chloroplasten, Amyloplasten, Mitochondrien, Vakuole, Cytosol oder Interzellularräume zu dirigieren, kann zwischen den Promotor und die kodierende Sequenz noch eine für eine sogenannte Signalsequenz oder ein Transitpeptid kodierende Sequenz gesetzt werden. In einigen Fällen ist es erforderlich, Sequenzen einzufügen, die für eine Signalsequenz am COOH-Terminus des Proteins kodieren. Die Sequenz muß im gleiPCT/EP99/07055 WO 00/17233 -5-

chen Leserahmen wie die kodierende Sequenz des Proteins sein. Zur Vorbereitung der Einführung der erfindungsgemäßen DNA-Sequenzen in höhere Pflanzen sind eine große Anzahl von Klonierungsvektoren verfügbar, die ein Replikationssignal für E.coli und einen Marker beinhalten, der eine Selektion der transformierten Zellen erlaubt. Je nach Einführungsmethode gewünschter Gene in die Pflanze können weitere DNA-Sequenzen erforderlich sein. Werden zum Beispiel für die Transformation der Pflanzenzelle das Ti- oder Ri-Plasmid verwendet, so muß mindestens eine rechte Begrenzung, häufig jedoch die rechte und die linke Begrenzung der Ti- und Ri-Plasmid T-DNA als Flankenbereich den einzuführenden Genen eingefügt werden. Die Verwendung von T-DNA für die Transformation von Pflanzenzellen ist intensiv untersucht und ausreichend in EP 120516; Hoekama, in: The Binary Plant Vector System, Offset-drukkerij Kanters B.V. Alblasserdam (1985), Chapter V; Fraley et al., Crit.Rev.Plant Sci. 4,1-46 und An et al. (1985) EMBO J. 4, 277-287 beschrieben worden. Ist die eingeführte DNA einmal im Genom integriert, so ist sie in der Regel stabil und bleibt auch in den Nachkommen der ursprünglich transformierten Zellen erhalten. Sie erhält normalerweise einen Selektionsmarker, der den transformierten Pflanzenzellen Resistenz gegenüber einem Biozid oder einem Antibiotikum, wie Kanamycin, G 418, Bleomycin, Hygromycin oder Phosphinotricin u.a. vermittelt. Der individuell verwendete Marker sollte daher die Selektion transformierter Zellen gegenüber Zellen, denen die eingefügte DNA fehlt, gestatten.

Für die Einführung von DNA in eine Pflanze stehen viele Techniken zur Verfügung. Diese Techniken umfassen die Transformation mit Hilfe von Agrobakterien, z.B. Agrobacterium tumefaciens, die Fusion von Protoplasten, die Mikroinjektion von DNA, die Elekroporation, sowie ballistische Methoden und die Virusinfektion. Aus dem transformierten Pflanzenmaterial können dann im geeigneten Medium, welches Antibiotika oder Biozide zur Selektion enthalten kann, wieder ganze Pflanzen regeneriert werden. Bei der Injektion und Elektroporation sind an sich keine speziellen Anforderungen an die Plasmide gestellt. Sollen aber aus derartig transformierten Zellen ganze Pflanzen regeneriert werden, ist die Anwesenheit eines selektierbaren Markergens not-

wendig. Die transformierten Zellen wachsen innerhalb der Pflanzen in der üblichen Weise (McCormick et al. (1986), Plant Cell Reports 5, 81-84). Die Pflanzen können normal angezogen werden und mit Pflanzen, die die gleiche transformierte Erbanlage oder andere Erbanlagen haben, gekreuzt werden. Die daraus entstehenden Individuen haben die entsprechenden phänotypischen Eigenschaften.

Weiterhin sind Gegenstand der Erfindung Expressionsvektoren, die eine oder mehrere der erfindungsgemäßen DNA-Sequenzen enthalten. Solche Expressionsvektoren erhält man, indem man die erfindungsgemäßen DNA-Sequenzen mit geeigneten funktionellen Regulationssignalen versieht. Solche Regulationssignale sind DNA-Sequenzen, die für die Expression verantwortlich sind, beispielsweise Promotoren, Operatoren, Enhancer, ribosomale Bindungsstellen, und die vom Wirtsorganismus erkannt werden.

Gegebenenfalls können noch weitere Regulationssignale, die beispielsweise Replikation oder Rekombination der rekombinanten DNA im Wirtsorganismus steuern, Bestandteil des Expressionsvektors sein.

Ebenso gehören die mit den erfindungsgemäßen DNA-Sequenzen oder Expressionsvektoren transformierten Wirtsorganismen zum Gegenstand der Erfindung.

Für die Expression der erfindungsgemäßen Enzyme eignen sich besonders solche Wirtszellen und Organismen, die keine intrinsischen Enzyme mit der Funktion der DOXP-Synthase, der DOXP-Reduktoisomerase oder des gcpE-Proteins aufweisen. Dies trifft für Archaebacterien, Tiere, Pilze, Schleimpilze und einige Eubakterien zu. Durch das Fehlen dieser intrinsischen Enzymaktivitäten wird die Detektion und Aufreinigung der rekombinanten Enzyme wesentlich erleichtert. Auch wird es erst dadurch möglich, mit geringem Aufwand die Aktivität und insbesondere die Hemmung der Aktivität der erfindungsgemäßen rekombinanten Enzyme durch verschiedenen Chemikalien und Pharmaka in Rohextrakten aus den Wirtszellen zu messen.

Die Expression der erfindungsgemäßen Enzyme erfolgt vorteilhafterweise dann in eukaryontischen Zellen, wenn posttranslatorische Modifikationen und eine native Faltung der Polypeptidkette erreicht werden soll. Außerdem wird in Abhängigkeit vom Expressionssystem bei der Expression genomischer DNA-Sequenzen erreicht, daß Introns durch Spleißen der DNA beseitigt und die Enzyme in der für die Parasiten charakteristischen Polypeptidsequenz produziert werden. Für Introns codierende Sequenzen können auch durch rekombinante DNA-Technologie aus den zu exprimierenden DNA-Sequenzen beseitigt oder experimentell eingefügt werden.

Die Isolierung des Proteins kann aus der Wirtszelle oder dem Kulturüberstand der Wirtszelle nach dem Fachmann bekannten Verfahren erfolgen. Es kann auch eine in vitro Reaktivierung der Enzyme erforderlich sein.

Zur Erleichterung der Aufreinigung können die erfindungsgemäßen Enzyme oder Teilsequenzen der Enzyme als Fusionsprotein mit verschiedenen Peptidketten exprimiert werden. Dazu eigenen sich besonders Oligo-Histidin-Sequenzen und Sequenzen, die von der Glutathion-S-Transferase, Thioredoxin oder Calmodulin-bindenden Peptiden abgeleitet sind.

Weiterhin können die erfindungsgemäßen Enzyme oder Teilsequenzen der Enzyme als Fusionsprotein mit solchen, dem Fachmann bekannten, Peptidketten exprimiert werden, daß die rekombinanten Enzyme in das extrazelluläre Millieu oder in bestimmte Kompartimente der Wirtszellen transportiert werden. Dadurch kann sowohl die Aufreinigung, als auch die Untersuchung der biologischen Aktivität der Enzyme erleichtert werden.

Bei der Expression der erfindungsgemäßen Enzyme kann es sich als zweckmäßig erweisen, einzelne Codone zu verändern. Dabei

WO 00/17233 -8-

ist der gezielte Austausch von Basen in der kodierenden Region auch sinnvoll, wenn die genutzten Codone in den Parasiten abweichend sind von der Codonnutzung im heterologen Expressionssystem, um eine optimale Synthese des Proteins zu gewährleisten.

Weiterhin können die erfindungsgemäßen Enzyme unter standardisierten Bedingungen durch dem Fachmann bekannte Techniken durch in vitro-Translation gewonnen werden. Dafür geeignete Systeme sind Kaninchen-Reticulozyten- und Weizenkeimextrakte und Bakterienlysate. Auch kann in vitro transskribierte mRNA in Xenopus-Oocyten translatiert werden.

Durch chemische Synthese können Oligo- und Polypeptide hergestellt werden, deren Sequenzen aus der Peptidsequenz der erfindungsgemäßen Enzyme abgeleitet sind. Bei geeigneter Wahl der Sequenzen besitzen derartige Peptide Eigenschaften, die für die erfindungsgemäßen Enzyme charakteristisch sind. Derartige Peptide können in großen Mengen hergestellt werden und eignen sich besonders für Studien über die Kinetik der Enzymaktivität, die Regulation der Enzymaktivität, die dreidimensionale Struktur der Enzyme, die Hemmung der Enzymaktivität durch verschiedenen Chemikalien und Pharmaka und die Bindungsgeometrie und Bindugnsaffinität verschiedener Liganden.

Vorzugsweise wird zur rekombinanten Herstellung der erfindungsgemäßen Enzyme eine DNA mit den Nukleotiden aus den Sequenzen SEQ ID NO: 1, 3 und 5 verwendet.

Die Erfindung umfaßt daher außerdem ein Verfahren zum Screening nach Verbindungen, die desDesoxy-D-xylulose-Phosphat-Stoffwechselweg inhibieren. Gemäß diesem Verfahren wird ein Wirtsorganismus, der einen rekombinanten Expressionsvektor enthält, wobei der Vektor zumindest einen Teil der Olignukleotidsequenz gemäß SEQ ID NO:1, SEQ ID NO: 3 oder SEQ ID NO: 5 oder Varianten oder Homologe dieser aufweist, und außerdem eine

Verbindung, von der vermutet wird, daß sie eine antimikrobielle, antiparasitare, antibakterielle, antivirale und antimykotische Wirkung bei Mensch und Tier oder eine antimikrobielle, antivirale, bakterizide, herbizide oder fungizide Wirkung bei Pflanzen hat, bereitgestellt. Anschließend wird der Wirtsorganismus mit der Verbindung in Kontakt gebracht und die Wirksamkeit der Verbindung bestimmt.

Ein weiterer Gegenstand dieser Erfindung sind Methoden zur Bestimmung der enzymatische Aktivität des gcpE-Proteins. Diese kann nach bekannten Verfahren bestimmt werden. Hierbei wird die Phosphorylierung eines Zuckers oder eines Phosphorzuckers oder einer Vorstufe der Isoprenoidbiosynthese, insbesondere die Phosphorylierung von 2-C-Methyl-D-erythritol, 2-C-Methyl-Derythritol-phosphat, insbesondere 2-C-Methyl-D-erythritol-4phosphat, 2-C-Methyl-D-erythrose, 2-C-Methyl-D-erythrosephosphat, insbesondere 2-C-Methyl-D-erythrose-4-phosphat, detektiert. Ein weiterer Gegenstand dieser Erfindung ist die Verwendung dieser Meßverfahren zur Ermittlung von Stoffen, die die Aktivität der jeweiligen Enzyme inhibieren.

Die enzymatische Aktivität von DOXP-Synthase und DOXP-Reduktisomerase kann in einem einzigen Schitt detektiert werden, indem die Umwandlung von Glycerinaldehyd-3-phosphat zu 2-C-Methylerythritol-4-phosphat bestimmt wird.

Analog erfolgt die Bestimmung der Aktivitäten von DOXP-Synthase und DOXP-Reduktoisomerase. Für die Bestimmung der DOXP-Synthase-Aktivität eignen sich auch fluorimetrische Verfahren, wie von Querol et al. beschrieben (Querol et al. Abstracts 4<sup>th</sup> european symposium on plant isoprenoids, Barcelona 21-23 April 1999).

### Patentansprüche

- DNA-Sequenzen, die für ein Polypeptid mit der in SEQ ID NO: 2 dargestellten Aminosäuresequenz codieren oder für ein Analoges oder Derivat des Polypeptids gemäß SEQ ID NO:2, worin eine oder mehrere Aminosäuren deletiert, hinzugefügt oder durch andere Aminosäuren substituiert worden sind.
- 2. DNA-Sequenzen, die für ein Polypeptid mit der in SEQ ID NO: 4 dargestellten Aminosäuresequenz codieren oder für ein Analoges oder Derivat des Polypeptids gemäß SEQ ID NO:4, worin eine oder mehrere Aminosäuren deletiert, hinzugefügt oder durch andere Aminosäuren substituiert worden sind.
- 3. DNA-Sequenzen, die für ein Polypeptid mit der in SEQ ID NO: 6 dargestellten Aminosäuresequenz codieren oder für ein Analoges oder Derivat des Polypeptids gemäß SEQ ID NO: 6, worin eine oder mehrere Aminosäuren deletiert, hinzugefügt oder durch andere Aminosäuren substituiert worden sind.
- 4. DNA-Sequenz gemäß einem der Ansprüche 1 bis 3, dadurch gekennzeichnet, daß sie außerdem funktionelle Regulationssignale, insbesondere Promotoren, Operatoren, Enhancer, ribosomale Bindungsstellen, aufweist.
- 5. DNA-Sequenz mit folgenden Teilsequenzen
  - i) Promotor, der in Viren, Eukaryonten und Prokaryonten aktiv ist und die Bildung einer RNA im vorgesehenen Zielgewebe oder den Zielzellen sicherstellt,
  - ii) DNA-Sequenzen gemäß einem der Ansprüche 1 bis 3,
  - iii) 3'-nichttranslatierte Sequenz, die in Viren, Eukaryonten und Prokaryonten zur Addition von Poly-A Resten an das 3'-Ende der RNA führt.
- 6. Verfahren zur Herstellung von transgenen Viren, Eukaryonten und Prokaryonten zur Veränderung des Isoprenoid-Gehaltes, dadurch gekennzeichnet, daß eine DNA-Sequenz gemäß Anspruch 4 oder 5 in das Genom von Viren, eukaryontischen und proka-

- ryontischen Zellen mit oder ohne Verwendung eines Vektors transferiert und eingebaut wird.
- 7. Transgene Systeme, insbesondere Pflanzen und Pflanzenzellen, welche ein oder mehrere DNA-Sequenzen gemäß der Ansprüche 1 bis 5 als "fremde" oder "zusätzliche" DNA enthalten, die exprimiert werden.
- 8. Expressionsvektor, enthaltend eine oder mehrere DNA-Sequenzen gemäß Anspruch 1 bis 5.
- 9. Protein, welches am 1-Deoxy-D-Xylulose-5-Phosphat-Stoffwechselweges beteiligt ist und a) codiert wird von den DNA-Sequenzen SEQ ID NO: 1,3 oder 5 oder b) codiert wird von DNA-Sequenzen, die mit den DNA-Sequenzen SEQ ID NO: 1,3,5 oder Fragmenten dieser DNA-Sequenzen im DNA-Bereich, der für das reife Protein codiert, hybridisieren.
- 10. Protein nach den Anspruch 9, erhältlich aus den Kulturüberständen von Parasiten oder aus den aufgeschlossenen Parasiten und Aufreinigung über chromatographische und elektrophoretische Techniken.
- 11. Protein nach einem der Ansprüche 9 und 10, dadurch gekennzeichnet, daß es a) das Produkt einer viralen, prokaryontischen oder eukaryontischen Expression einer exogenen DNA ist, b) codiert wird von den Sequenzen SEQ ID NO: 1, 3 oder 5 oder codiert wird von DNA-Sequenzen, die mit den in den DNA-Sequenzen SEQ ID NO: 1, 3 oder 5 oder Fragmenten dieser DNA-Sequenzen im DNA-Bereich, der für das reife Protein kodiert, hybridisieren, oder c) codiert wird von DNA-Sequenzen, die ohne Degeneration des genetischen Codes mit den in b) definierten Sequenzen hybridisieren würden und für ein Polypeptid mit entsprechender Aminosäure-Sequenz kodieren.

- 12. Protein gemäß einem der vorangehenden Ansprüchen, dadurch gekennzeichnet, daß es die Aminosäuresequenzen SEQ ID NO: 2, 4 oder 6 aufweist.
- 13. Verfahren zur Bestimmung der enzymatischen Aktivität des gcpE-Proteins, dadurch gekennzeichnet, daß Phosphorylierung eines Zuckers oder eines Phosphorzuckers oder einer Vorstufe der Isoprenoidbiosynthese, insbesondere die Phosphorylierung von 2-C-Methyl-D-erythritol, 2-C-Methyl-D-erythritol-erythritol-phosphat, insbesondere 2-C-Methyl-D-erythritol-4-phosphat, 2-C-Methyl-D-erythrose, 2-C-Methyl-D-erythrosephosphat, insbesondere 2-C-Methyl-D-erythrose-4-phosphat, und der Phosphat- und Alkoholvorstufen, detektiert wird.
- 14. Verfahren nach Anspruch 13, dadurch gekennzeichnet, daß die Phosphorylierung der folgenden Phosphate oder Alkohole detektiert wird:

 $CH_2(OH) - C(CH_3) = C(OH) - CH_2 - O - PO(OH)_2$ 

 $CH_2 (OH) - C (CH_3) = C (OH) - CH_2 - OH$ ,

 $CH_{2}$  (OH) -CH (CH<sub>3</sub>)  $-CO-CH_{2}-O-PO$  (OH)  $_{2}$ ,  $CH_{2}$  (OH) -CH (CH<sub>3</sub>)  $-CO-CH_{2}-OH$ 

 $CH_2=C (CH_3) -CO-CH_2-O-PO (OH)_2$ ,  $CH_2=C (CH_3) -CO-CH_2-OH$ ,

 $CH_2=C (CH_3) - CH (OH) - CH_2 - O - PO (OH)_2$ ,  $CH_2=C (CH_3) - CH (OH) - CH_2 - OH$ ,

 $CH_2$  (OH) -C (= $CH_2$ ) -C (OH) - $CH_2$ -O-PO (OH) 2,

 $CH_{2}$  (OH) -C (= $CH_{2}$ ) -C (OH) - $CH_{2}$ -OH

 $\label{eq:cho-ch} \text{CHO-CH (CH}_3) - \text{CH (OH)} - \text{CH}_2 - \text{O-PO (OH)}_2, \quad \text{CHO-CH (CH}_3) - \text{CH (OH)} - \text{CH}_2 - \text{OH},$ 

 $CH_2(OH) - C(OH)(CH_3) - CH = CH - O - PO(OH)_2$ ,

 $CH_2$  (OH) -C (OH) (CH<sub>3</sub>) -CH=CH-OH

 $CH(OH) = C(CH_3) - CH(OH) - CH_2 - O - PO(OH)_2$ ,

CH (OH) =  $C(CH_3)$  - CH(OH) -  $CH_2$ - OH,

 $(CH_3)_2HC-CO-CH_2-O-PO(OH)_2$ ,

 $(CH_3)_2HC-CO-CH_2-O-H_2$ 

 $(CH_3)_2HC-CH(OH)-CH_2-O-PO(OH)_2$ ,

(CH<sub>3</sub>)<sub>2</sub>HC-CH (OH) -CH<sub>2</sub>-O-H.

15. Verfahren zur gekoppelten Bestimmung der enzymatischen Aktivität der DOXP-Synthase und der DOXP-Reduktase, dadurch gekennzeichnet, daß die Umwandlung von Glycerinaldehyd-3-phosphat zu 2-C-Methylerythritol-4-phosphat detektiert wird.

- 16. Verfahren zum Screening einer Verbindung für die Therapie von infektiösen Prozessen bei Mensch und Tier, wobei das Verfahren umfaßt:
  - a) Bereitstellen einer Wirtszelle, die einen rekombinanten Expressionsvektor enthält, wobei der Vektor zumindest einen Teil der Olignukleotidsequenz gemäß SEQ ID NO:1, SEQ ID NO: 3 oder SEQ ID NO: 5 oder Varianten oder Analoga dieser aufweist, und außerdem eine Verbindung, von der vermutet wird, daß sie eine antimykotische, antibiotische, antiparasitäre oder antivirale Wirkung bei Mensch und Tier hat,
  - b) In-Kontakt-Bringen der Wirtszelle mit der Verbindung und
  - c) Bestimmung der antimikrobiellen, antimykotischen, antibiotischen, antiparasitären oder antiviralen Wirksamkeit der Verbindung.
- 17. Verfahren zum Screening nach Verbindungen zur Behandlung von Pflanzen, wobei das Verfahren umfaßt:
  - a) Bereitstellen einer Wirtszelle, die einen rekombinanten Expressionsvektor enthält, wobei der Vektor zumindest einen Teil der Olignukleotidsequenz gemäß SEQ ID NO:1, SEQ ID NO: 3 oder SEQ ID NO: 5 oder Varianten oder Analoga dieser aufweist, und außerdem eine Verbindung, von der vermutet wird, daß sie eine antimikrobielle, antivirale, antiparasitäre, bakterizide, fungizide oder herbizide Wirkung bei Pflanzen hat,
  - b) In-Kontakt-Bringen der Wirtszelle mit der Verbindung und
  - c) Bestimmung der antimikrobiellen, antiviralen, antiparasitären, bakteriziden, fungiziden oder herbiziden Wirksamkeit der Verbindung.
- 18. Verwendung von DNA nach einem der Ansprüche 1 bis 5 oder von Proteinen nach einem der Ansprüche 9 bis 12 oder von transgenen Systemen nach Anspruch 7 zur Vorbeugung oder Therapie von Erkrankungen bei Mensch und Tier.

```
<110> Jomaa, Hassan
<120> Gene des 1-Desoxy-D-xylulose-Biosynthesewegs
<130> 15696
<140> PCT/EP99
<141> 1999-09-22
<150> DE19923567.8
<151> 1999-05-22
<150> DE19843279.8
<151> 1998-09-22
<160> 6
<170> PatentIn Ver. 2.1
<210> 1
<211> 1467
<212> DNA
<213> Plasmodium falciparum
<220>
 <221> CDS
<222> (1)..(1467)
 <220>
 <221> gene
 <222> (1)..(1467)
 <220>
 <221> mRNA
 <222> (1)..(1467)
 <400> 1 .
 Met Lys Lys Tyr Ile Tyr Ile Tyr Phe Phe Ile Thr Ile Thr Ile
```

10

															aga -	96
Asn	Asp	Leu	Val	Ile	Asn	Asn	Thr	Ser	Lys	Cys	Val	Ser	Ile	Glu	Arg	
			20					25					30			
•												•				
								tat								144
Arg	Lys	Asn	Asn	Ala	Tyr	Ile	Asn	Tyr	Gly	Ile	Gly	Tyr	Asn	Gly	Pro	
		35					40					45				
							•		•							
								aga								192
Asp	Asn	Lys	Ile	Thr	Lys	Ser	Arg	Arg	Cys_	Lys	Arg	Ile	Lys	Leu	Cys	
	50					55					. 60					
								,			•. •					
								gca								240
Lys	Lys	Asp	Leu	Ile	Asp	Ile	Gly	Ala	Ile		Lys	Pro	Ile	Asn		
65					70	·				75					80	
												.:				200
								ata								288
Ala	Ile	Phe	Gly	Ser	Thr	Gly	Ser	Ile		Thr	Asn	Ala	Leu		116	
				85					90					95		
																226
								aat								336
Ile	Arg	Glu		••	Lys	Ile	Glu	Asn	Val	Phe	Asn	vaı			ren	•
			100	1				105					110			
															***	384
								tta -								304
Tyr	Val	•		Ser	Val	Asn		Leu	Tyr	GIU	GIN			GIU	rne	
		115	•				120					125				
								4-	٠.			+-+	~~~	<b>~</b> ~~	++=	432
								gat								132
Leu			туг	Leu	Cys			Asp	гåз	Şer	140		GIU	GIU	beu	
	130	)				135	)				140					
									4-4			2+2	2+2	++4	tat	480
								gat								400
		ı Leı	ı Val	LLys			: Lys	Asp	Tyr			TTE	116	Leu	160	
145					150	)			•	155	)				100	
									4				<b>.</b> .		222	528
								tgt								320
Gly	/ Ası	Gl:	u Gl			s Glu	ı Ile	Cys			ASN	ser	116			
				169	5				170	,				175		
																<b>57</b> 6
ats	01	t at	- 00	t ati	- gai	t tct	ttt	. caa	· qqa	ı tta	ı tat	tct	act	: atg	tat	576

Ile	Val	Ile	Gl	.y 1	Ile	Asp	Ser	Phe	Gln	Gly	Leu	Tyr	Ser	Thr	Met	Tyr	
	٠		18						185	•	•			190		•	
			•												•		
σcà	att	ato	ı aa	at a	aat	aaa	ata	gtt	gcg	tta	gct	aat	aaa	gaa	tcc	att	624 ·
Ala	Ile	Met	A.	sn i	Asn	Lys	Ile	Val	Ala	Leu	Ala	Asn	Lys	Glu	Ser	Ile	
		195				•		200				•	205	•	٠	•	
	+ 0+	act	- ~	~+	ttc	+++	tta	ааσ	aaa	tta	tta	aat	att	cat	aaa	aat	672
y.L	C	. gc		y c 1	Dhe	Pha	Len	Lvs	Lvs	Leu	Leu	Asn	Ile	His	Lys	Asn	
var			ı G.	гy	FIIE	1110	215	_,_	-1-			220					
	210						213									•	
			_				<b></b> +		<b>722</b>	cat	ant	act	ata	ttt	caa	tat	720
.gca	aag	ata	a a	ta	CCT	gu.	gat	Com	C1	uie	Sor	.gcc.	Tle	Phe	Gln	Cvs	
	Lys	110	e I	1e	Pro		Asp	Ser	GIU	nra			Ile			240	
225						230	•				235	•				210	
														~~~	22+	+++	768
tta	gat	aa	t a	at	aag	gta	tta	aaa	aca	aaa	tgt -	tta	caa	gac	2	Dho	
Leu	Ası	As	n A	sn	Lys	Val	Leu	Lys	Thr		Cys	Leu	Gln	Asp		File	•
					245				·	250					255		
•									•		•						016 -
tct	aaa	a at	t a	ac	aat	ata	aat	aaa	ata	ttt	tta	tgt	tca	tct	gga	ggt.	816
Ser	Ly	s Il	e P	lsn	Asn	Ile	Asn	Lys	Ile	Phe	Leu	Cys	Ser			g1À.	
	٠		2	260	•				265			••	•	270	ı		
													•	•			
CC	tt	t ca	a a	aat	tta	act	atg	gad	gaa	tta	aaa	aat	gta	aca	tca	gaa	864
Pro	Ph	e Gl	n A	Asn	Leu	Thr	Met	Asp	Glu	Leu	Lys	Asn	Val	Thr	Ser	Glu	
٠.		27						280					285				
												·		•			
aat	t gc	t tt	a a	aag	cat	cct	aaa	a tgg	aaa	atg	ggt	aag	, aaa	ata	act	ata	912
Ası	n Al	a Le	eu I	Lys	His	Pro	Lys	Tr	Lys	Met	Gly	Lys	Lys	Ile	Thi	Ile	
	29			-			295					300					
																	•
αa	r to	t a	ca	act	ato	ato	r aat	t aaa	a ggt	: tta	gag	ggtt	: ata	gaa	acc	cat	960
y.c	n Sc	r A	1 =	Thr	Met	. Met	Ası	n Lv:	s Gİv	, Let	Gli	ı Val	l Ile	Glu	ı Thi	His	
30						310			-		319					320	
30	<b>J</b>					J.,	•										
									F 421	r sts	a mai	a ofti	r ata	ata	a cat	aaa	1008
tt	t t	a t	tt	gat	gta	a ga	- ta	C dd	L gai	. Tl	. C1		1 112	. Va	1 41	aaa Lvs	
Ph	e Le	eu P	he	Asp			р Ту	r Asi	n AS)			u va.	F T76	, va.	33	s Lys	
		•			325	5				330	J				.در	•	
•									•								1056
ga	a't	gc a	ţţ	ata	a ca	t tc	t tg	t gt	t ga	a tt	t at	a ga	c aaa	a tc	a gt	a ata	1036
G1	u C	ys I	le,	Ile	e Hi	s Se	r Cy	s Va	1 G1	u Ph	e Il	e As	p Ly:			l Ile	
				340	0				34	5				35	0		

agt	caa	atg	tat	tat	cca	gat	atg	caa	ata	ccc	ata	·tta	tat	·tct	tta	1104
Ser	Gln	Met	Tyr	Tyr	Pro	Asp	Met	Gln	Ile	Pro	Ile	Leu	.Tyr	Ser	Leu	
		355					360					365				•
aca	tgg	cct	gat	aga	ata	aaa	aca	aat	tta	aaa	cct	tta	gat	ttg	gct	1152
			_	Arg												
	370		_	_		375					380				,	•
cag	gtt	tca	act	ctt	aca	ttt	cat	aaa	cct	tct	tta	gaa	cat	ttc	ccg	1200
_	-			Leu												
285					390			_		395					400	
tat	att	aaa	tta	gct	tat	caa	qca	ggt	ata	aaa	qqa	aac	ttt	tat	cca	1248
-				Ala												
		-,-		405	- 4				410	•	•			415		
						•										
act	σta	ста	aat	gcg	tca	aat	gaa	ata	act	aac	aac	tta	ttt	tta	aat	1296
				Ala												
		202	420		•••			425					430			
aat	aaa	att	222	tat	ttt	gat	att	tcc	tct	ata	ata	tca	caa	att	ctt	1344
				Tyr												
	-,-	435	-,0	24-			440					445				
		133					•••									
gaa	FCt	ttc	, aat	tct	caa	aaq	att	t.ca	gaa	aat	agt	gaa	gat	tta	ato	1392
				Ser												
Giu	450		Noil	Ser	GIII	455	*41	501	014		460	O.L.	p	200		
	430					433					400	•				
		- 4- 4-	- •													1440
				caa												1440
	GIn	lle	Leu	Gln		His	ser	Trp	Ala		Asp	гуs	Ala	Thr		•
465					470					475					480	
				cat				tag								1467
Ile	Tyr	Asn	Lys	His		Ser	Ser						٠.			
				485												

<210> 2

<211> 488

<212> PRT

. <213> Plasmodium falciparum

e 1	00>	2

Met Lys Lys Tyr Ile Tyr Ile Tyr Phe Phe Phe Ile Thr Ile Thr Ile

1 5 10 15

Asn Asp Leu Val Ile Asn Asn Thr Ser Lys Cys Val Ser Ile Glu Arg 20 25 30

Arg Lys Asn Asn Ala Tyr Ile Asn Tyr Gly Ile Gly Tyr Asn Gly Pro

Asp Asn Lys Ile Thr Lys Ser Arg Arg Cys Lys Arg Ile Lys Leu Cys 50 55 60

Lys Lys Asp Leu Ile Asp Ile Gly Ala Ile Lys Lys Pro Ile Asn Val 65 70 75 80

Ala Ile Phe Gly Ser Thr Gly Ser Ile Gly Thr Asn Ala Leu Asn Ile 85 90 95

Ile Arg Glu Cys Asn Lys Ile Glu Asn Val Phe Asn Val Lys Ala Leu 100 105 110

Tyr Val Asn Lys Ser Val Asn Glu Leu Tyr Glu Gln Ala Arg Glu Phe 115 120 125

Leu Pro Glu Tyr Leu Cys Ile His Asp Lys Ser Val Tyr Glu Glu Leu 130 135 140

Gly Asp Glu Gly Met Lys Glu Ile Cys Ser Ser Asn Ser Ile Asp Lys 165 170 175

Ile Val Ile Gly Ile Asp Ser Phe Gln Gly Leu Tyr Ser Thr Met Tyr 180 185 190

Ala Ile Met Asn Asn Lys Ile Val Ala Leu Ala Asn Lys Glu Ser Ile 195 200 205

Val Ser Ala Gly Phe Phe Leu Lys Lys Leu Leu Asn Ile His Lys Asn

220

210

Ala Lys Ile Ile Pro Val Asp Ser Glu His Ser Ala Ile Phe Gln Cys 225 230 235 240

215

Leu Asp Asn Asn Lys Val Leu Lys Thr Lys Cys Leu Gln Asp Asn Phe
245 250 255

Ser Lys Ile Asn Asn Ile Asn Lys Ile Phe Leu Cys Ser Ser Gly Gly 260 265 270

Pro Phe Gln Asn Leu Thr Met Asp Glu Leu Lys Asn Val Thr Ser Glu 275 280 285

Asn Ala Leu Lys His Pro Lys Trp Lys Met Gly Lys Lys Ile Thr Ile 290 295 300

Asp Ser Ala Thr Met Met Asn Lys Gly Leu Glu Val Ile Glu Thr His 305 310 315 320

Phe Leu Phe Asp Val Asp Tyr Asn Asp Ile Glu Val Ile Val His Lys
325 330 335

Glu Cys Ile Ile His Ser Cys Val Glu Phe Ile Asp Lys Ser Val Ile 340 345 350

Ser Gln Met Tyr Tyr Pro Asp Met Gln Ile Pro Ile Leu Tyr Ser Leu 355 360 365

Thr Trp Pro Asp Arg Ile Lys Thr Asn Leu Lys Pro Leu Asp Leu Ala 370 375 380

Gln Val Ser Thr Leu Thr Phe His Lys Pro Ser Leu Glu His Phe Pro 385 390 395 400

Cys Ile Lys Leu Ala Tyr Gln Ala Gly Ile Lys Gly Asn Phe Tyr Pro 405 410 415

Thr Val Leu Asn Ala Ser Asn Glu Ile Ala Asn Asn Leu Phe Leu Asn 420 425 430

Asn Lys Ile Lys Tyr Phe Asp Ile Ser Ser Ile Ile Ser Gln Val Leu

435

445

Glu Ser Phe Asn Ser Gln Lys Val Ser Glu Asn Ser Glu Asp Leu Met 450 455 460

440

Lys Gln Ile Leu Gln Ile His Ser Trp Ala Lys Asp Lys Ala Thr Asp 465 470 475 480

Ile Tyr Asn Lys His Asn Ser Ser 485

<210> 3

<211> 3872

<212> DNA

<213> Plasmodium falciparum

<220>

<221> CDS

<222> (126)..(3740)

<220>

<221> gene

<222> (1) .. (3870)

<220>

<221> mRNA

<222> (1)..(3870)

<400> 3

ggtaatatac gtataatata tatataatat attcttacgt atgtatcatt tatgaatcat 60

aataatatto taaatttaco ttoogttttt gotogatott otoattttog tttoagottt 120

tatca atg att ttt aat tat gtg ttt ttt aag aac ttt gta cca gtt gtt 170 Met Ile Phe Asn Tyr Val Phe Phe Lys Asn Phe Val Pro Val Val

5 10 15

25

cta tac att ctc ctt ata ata tat att aac tta aat ggc atg aat aat 218 Leu Tyr Ile Leu Leu Ile Ile Tyr Ile Asn Leu Asn Gly Met Asn Asn

.

20

30

									•	٠.						
			ata													266
Lys	Asn	Gln	Ile	Lys	Thr	Glu	Lys	Ile	Tyr	Ile	Lys	Lys	Leu	Asn	Arg	
			35					40				,	45			
			aaa													314
Leu	Ser		Lys	Asn	Ser	Leu		Ser	Ser	Lys	Asn		Ile	Ala	Cys	
		50					55					60				
<b></b>		,					~~+	224	3.03	22+	200	202	tat	000	tat	362
			ata Ile													302
Leu	65	ASP	116	GIY	ASII	·70	ush	7.011	<i></i> 4	71311	75			<b>4-</b> 3	-1-	
	03															
aat	ata	aat	gtt	aaa	aat	gat	gat	att	aat	tcc	tta	cta	aaa	aat	aat	410
			Val													
-80				-	85	_				90			•		95	
						·										
tat	agt	aat	aaa	ttg	tac	atg	gat	aag	agg	aaa	aat	att	aat	aat	gta	458
Tyr	Ser	Asn	Lys	Leu	Tyr	Met	Asp	Lys	Arg	Lys	Asn	Ile	Asn	Asn	Val	
			•	100					105					110		
			aat													506
Ile	Ser	Thr	Asn	Lys	Ile	Ser	Gly		Ile	Ser	Asn	Ile		Ser	Arg	
			115					120					125			
														***	200	554
•			gaa													334
ASN	GIN	•	Glu	Asn	GIU	GIU	135	AIG	ASII	гур	GIII	140		Leu	1111	
		130					133					110				
caa	tat	cac	act	tat	aat	atq	tca	cat	gaa	caσ	αac	aaa	cta	act	aat	602
			Thr													
	145			- , -		150					155	•				
gat	aat	aat	agg	aat	aat	aaa	aag	aat	ttt	aat	tta	tta	ttt	ata	aat	650
			Arg													
160					165				•	170					175	
tat	ttt	. aat	ttg	aaa	cga	atg	aaa	aat	tct	ctt	cta	aat	aaa	gac	aat	698
Tyr	Phe	Asr	Leu	Lys	Arg	Met	Lys	Asn	Ser	Leu	Leu	Asn	Lys	Asp	Asn	
				180					185					190		
tro		tac	: tat		gaa	aaa	aaa	tta	tca	ttt	cta	cat	aaq	gcc	tat	746

									,							
Phe	Phe	Tvr	Cvs	Lvs	Glu	Lys	Lys	Leu	Ser	Phe	Leu	His	Lys	Ala	Tyr	
			195			_		200					205			
			193											•		
											**-		202	222	tct	794
					act											
Lys	Lys	Lys	Asn	Cys	Thr	Phe	Gln	Asn	Tyr	Ser	Leu	Lys	Arg	rys	Ser	•
		210					215					220				
2 2 t	cat	ast	tca	cat	aaa	tta	ttt	tct	gga	gaa	ttt	gac	gat	tat	aca	842
					Lys											
Asn			Ser	ura	цуз				,		235		•			
	225					230					233					•
														•		
					tat											890
Asn	Asn	Asn	Ala	Leu	Tyr	Glu	Ser	Glu	Lys	Lys	Glu	Tyr	Ile	Thr	Leu	
240					245					250					255	
														•		
					aat	aat	aat	aat	aaa	aat	aat	gat	aat	aaa	aat	938
					Asr											
Asn	Asn	ASI	1 ASI			i Maii	, ASII	ASII		11011				270		
				260	)				265					2.0		
															•	006
					t tat											986
Asn	Asp	As:	n As	n As	р Туг	. Asr	Asn	Asn	Asn	Ser	Cys	Asn	Asn	Leu	Gly	
			27					280					285			
				<b>.</b>	t ta	r dat	- aat	tat	gat	aga	gat	aat	aat	aat	сса	1034
Glı	ı Ar	g Se	r As	u HI	s Ty	r Asi			. Gry	GLy	, wah				Pro	
		29	0				295	•				300				
tgi	t aa	t aa	t aa	t aa	t ga	c aa	a tai	t gat	ata:	gga	a aaa	tat	ttc	: aaa	cag	1082
															Gln	
•	30					31					315					
	50	•				,										
							<b>.</b>	+ ~~	. + = +	:	a act	- ata	tat	aat	gat	1130
															gat Asn	•
11	e As	n Th	ır Pl	ne Il	Le As	n Il	e As	p GI	u Tyı			c 11e	ı ıyı	L GI	Asp	
32	0				32	!5	1			33	0				335	•
αa	a at	a ta	at a	aa o	aa at	a ta	t ga	a ct	a tai	gt	a gaa	a aga	aat	t at	cct	1178
															e Pro	•
G.I.	u II	LE I	Ar r						34		2.2		-	350		
				3	40				343	•				23	-	
																1000
ga	a ta	at t	at g	aa c	ga aa	aa ta	t tt	t to	a ga	a ga	t at	t aaa	a aa	g ag	t gtc	1226
G1	u T	yr T	yr G	lu A	rg L	ys Ty	r Ph	e Se	r Gl	u As	p Il	e Ly	s Ly	s Se	r Val	
	•	-		55				36					36			•
			-													

С	ta	ttt	gat	ata	gat	aaa	tat	aat	gat	gtc	gaa	ttt	gaa	aaa	gct	ata	1274
L	eu	Phe	Asp	Ile	Asp	Lys	Tyr	Asn	Asp	Val	Glu	Phe	Glu	Lys	Ala	Ile	
			370					375			٠	٠	380				7
	•																
a	aa	gaa	gaa	ttt	ata	aat	aat	gga	gtt	tat	att	aát	aat	ata	gat	aat.	1322
										Tyr							•
		385				•	390					395					
ā	ica	tat	tat	aaa	aaa	gaa	aat	att	tta	ata	atg	aaa	aag	ata	tta	cat	1370
										Ile							
	100	- 4 -	- 4 -		-3-	405					410					415	•
												٠.					
ŧ	аг	ttc	cca	tta	tta	222	tta	att	aat	aat	cca	tca	gat	tta	aaa	aag	1418
										Asn							
•	. <u>y</u> -	- 110	110	Deu	420	2,0	200			425			•		430	•	
					420												
		222	222		+=+	++=	cct	tta	tta	gca	cat	gaa	tta	·aaa	ata	ttt	1466
										Ala							
•	beu	Буз	БÃЗ	435	1 7 1	Dea		<b></b>	440					445			•
				433					•••								
							-+-	202	aa2	ggt	cat	+++	t.cc	tct	att	tta	1514
										Gly							
	Leu	Pne			vai	Asn	TIE		GIY	GIY	nis	FILE	460	Jer	<b>V</b> 41	Dea	•
			450					455					400				
																003	1562
										ttg							1302
	Ser			Glu	Ile	Gln			Leu	Leu	Tyr		Pne	ASII	GIII	FIO	
		465	1				470					475					
																	1610
										cat							1610
	Tyr	Asp	Asr	Val	Ile	Tyr	Asp	Ile	Gly	His			Tyr	Val	His		
	480				•	485	•				490					495	
																٠.	
	ata	ttç	aco	gga	a aga	aaa	cta	tta	ttt	cta	tca	tta	aga	aat	aaa	aaa	1658
	Ile	Leu	Th:	Gly	y Arc	Lys	Let	Leu	Phe	Leu	Ser	Leu	Arg	Asn	Lys	Lys	
					500	)				505					510		
	ggt	att	agi	t gga	a tto	cta	a aat	att	ttt	gaa	agt	att	tat	gat	aaa	ttt	1706
										Glu							
	•			51					520					525			
	aaa	ו מכי	ר ממ	t car	c 201	t to	c act	t tca	a tta	agt	gct	ata	caa	gga	tat	tat	1754
	744	, y~	- 44		- 44		,			_	-						

(	Gly	Ala	Gly	His	Ser	Ser	Thr	Ser	Leu	Ser	Ala	Ile	Gln	Glý	Tyr	Tyr	
			530					535					540		٠.	•	
					.•			Δ.				<b>.</b> . <b>.</b>					1802
						gtg											1002
	Glu			Trp	Gln	Val		Asn	гЛя	GIU	гàг	555	GIY	ASII	GLY	vab	
		545					550					333					
	ata	gaa.	ata	agt	gat	aac	qca	aat	gtc	acg	aat	aat	gaa	agg	ata	ttt	1850
						Asn											
	560				•	565					570					575	
						aat											1898
	Gln	Lys	Gly	Ile	His	Asn	Asp	Asn	Asn	Ile	Asn	Asn'	Asn	Ile	Asn	Asn	
					580					585					590		
																	1946
																aat	1940
	Asn	Asn	Tyr	595		Pro	Ser	ASP	600	AGI	GIY	ALG	Giu	605	1111	11011	
				333	,												
	qta	сса	aat	gta	cga	aat	gat	aac	cat	aac	gtg	gat	aaa	gta	cac	att	1994
						Asn											
			610	)				615					620				
						ggt											2042
	Ala	Ile	e Ile	e Gl	y Asp	Gly			Thr	Gly	Gly			Leu	Glu	Ala	
		625	5				630					635					
			<b>.</b>			a ttc			+ < +	222	att	tta	att	att	tat	aat	2090
						r Phe											
	640		ııy.	1 11	. 361	645					650		,		•	655	
						-											•
	gat	aa	c gg	a ca	a gti	t tct	tta	cca	aca	aat	gcc	gta	agt	ata	tca	ggt	2138
						l Ser											
					66	0				665	5				670	)	
					٠.												
						t tct											2186
	Ası	n Ar	g Pr	o I1	e Gl	y Sei	: Ile	Se Se			s Lev	ı His	Туг			Ser	•
				67	5				680	)				685	· ·		
											- ++-		T 227				2234
																a aaa a Lvs	
	AS	n 11	e G1. 69		a As	U AT	a GT	9 ASI 69:		. ny:	ວ ກະເ	- 56	700			a Lys	
			0:	, ,					_								

									T	Z				•		
			att													2282
Glu	Asņ	Asn	Ile	Phe	Glu	Asn	Leu	Asn	Tyr	Asp	Tyr	Iļe	Gly	Val	Val	
	705					710					715				·	
aat	ggt	aat.	aat	aca	gaa	gag	ctc	ttt	aaa	gta	tta	aat	aat	ata	aaa.	2330
Asn	Gly	Asn	Asn	Thr	Glu	Glu	Leu	Phe	Lys	Val	Leu	Asn	Asn	Ile	Lys	
720					725					730					<b>735</b> .	
gaa	aat	aaa	tta	aaa	aga	gct	act	gtt	ctt	cat	gta	cgt	aca	aạa	aaa	2378
Glu	Asn	Lys	Leu	Lys	Arg	Ala	Thr	Val	Leu	His	Val	Arg	Thr	Lys	Lys	•
				740					745					750		•
											٠.					
tcg	aat	gat	ttt	ata	aat	tca	aag	agt	сса	ata	agt	ata	ttg	cac	tct	2426
			Phe											-		
		-	755					760					765			
ata	aag	aaa	aat	gag	att	ttc	cct	ttc	gat	acc	act	ata	tta	aat	gga	2474
			Asn													
	•	770					775					780				
aat	att	cat	aag	gag	aac	aaq	ata	gaa	gaa	gag	aaa	aat	gtg	tct	tca	2522
			Lys													
	785		-1-	-	•	790					795					•
tet	aca		tat	αat	ata	aat	aat	aag	aat	aat	aaa	aat	aat	qat	aat,	2570
			Tyr													
800		,_	-,-		805		;-	-		810	-				815	
					000											
ant	<b>~</b> 3=		ata		. +=+	a a a	gat	atα	ttt	tca	222	gag	acq	ttc	aca	2618
			: Ile													
Ser	GIL	1 116	: 116			GIU	nsp	1100	825	001	2,0	020		830		
				820	•				Ų23					050		
									+-+	++-	226	222	ast	242	227	2666
			aca													2000
Asp	116	e Ty	Thr		GIU	Met	Leu			Leu	гÀг	гàг		ALG	ASII	
			835	j				840	٠				845			
																071.
			c cta													2714
Ile	110	e Ph	e Lev	ı Sez	Pro	Ala			Gly	Gly	Ser			Val	Lys	
		85	0				855	,				860	٠.			
art	aσ	t ga	a cat	t tai	cca	aat	aat	gta	tat	gat	gta	ggt	ata	gca	gaa	2762

	_		_	_	_	<b>&gt;</b>	X	Va1	T1/~	Acn	V=1	Glv	Tle	Ala	Glu	
Ile	Ser	Glu	Arg	Tyr			ASII	Val	TYL	rap		Gly				
	865					870					875	•	•	•		
		•														
caa	cat	tct	gta	act	ttc	gca	gca	gct	atg	gca	atġ	aat	aag	aaa	tta	2810
												Asn				
					885					890					895	
880					883					•••						
												aga				2858
Lys	Ile	Gln	Leu	Суз	Ile	Tyr	Ser	Thr	Phe	Leu	Gln	Arg	Ala	Tyr	Asp	
				900					905					910		
								•								
<b>633</b>	2++	a t- a	cat	ast	ctt	aat	tta	caa	aat	ata	cct	tta	aag	gtt	ata	2906
												Leu				
GIn	He	TTe			rea	ASI	Leu		nsii	116		, mea			:	•
٠			915					920					925			
																•
att	gga	aga	agt	gga	tta	gta	gga	gag	gat	ggg	gca	aca	cat	caa	ggt	2954
Ile	Glv	Arc	Ser	Glv	Leu	Val	Gly	Glu	Asp	Gly	Ala	Thr	His	Gln	Gly	
	•	930		_			935				•	940				
		JŸC					•••									
									a++	220	221		tat	ata	ata	3002
												gca				
Ile	Ту	Ası	Lei	ı Ser	Tyr	Leu	Gly	Thr	Leu	Asn		Ala	Tyr	116	116	
	945	5				950					955	5				
tct	cca	a ag	t aa	t caa	att	gat	ttg	aaa	aga	gct	cti	agg	ttt	gct	tat	3050
												ı Arg				
		J 36.	L Awai	11 (31)				- <b>-</b> -		970		_			975	
960	).				965	•				370					•	
			٠				•									2000
															ata	3098
Le	u As	р Lу	s As	p Hi	s Sea	c Val	. Tyı	r Ile	Arq	, Ile	Pr	o Arg	Met	Asn	Ile	
				986					985					990		
						•										
								a +a1	- tt	7 220	at	t cat	ato	ı aaa	aat	3146
Le	u Se	r As	p Ly	s Ty	r Me	t Ly:	s GI	у ту	r Lei	ı Ası	1 11	e nis			a Asn	
			99	15				100	0				1005	)		
													,			
σa	σ ac	ic aa	a aa	t at	c ga	t ata	a aa	c gt	g ga	t ata	a aa	c gat	gat	gta	a gat	3194
22	, -,		n-	n T1	- 3-	n Va	l As	n Va	 1 Ası	p Il	e As	n Ası	Ası	va!	l Asp	
GI	u se			):: II	e no	P 44						102			_	•
		- 101	10				101	J				1021	-			
•																2010
aa	a ta	at a	gt ga	aa ga	a ta	t at	g ga	c ga	t ga	t aa	t tt	t at	a aaa	a tc	g ttt	3242
Ly	s T	r S	er G	lu Gl	u Ty	r Me	t As	p As	p As	p As	n Ph	e Il	e Ly	s Se	r Phe	
•	10					103					103					•

			tct													3290
le (	Gly	Lys	Ser	Arg	Ile	Ile	Lys	Met	Asp	Asn	Glu	Asn	Asn	Asn	Thr	
1040				1	045				3	1050	•			1	055	
•																
aat	gaa	cat	tat	tca	agc	aga	gga	gat	aca	cag	aca	aaa	aaa	aaa	aaa	3338
Aşn (	Glu	His	Tyr	Ser	Ser	Arg	Gly	Asp	Thr	Gln	Thr	Lys	Lys	Lys	Lys	
			1	1060				1	1065					1070		
:																
gtt	tgt	atc	ttt	aac	atg	ggt	agt	atg	ctt	ttt	aat	gta	att	aat	gct	3386
Val	Cys	Ile	Phe	Asn	Met	Gly	Ser	Met	Leu	Phe	Asn	Val	Ile	Asn	Ala	
		;	1075				:	1080				. :	1085			•
											٠					
ata	aaa	gaa	att	gaa	aaa	gaa	caa	tat	att	tca	cat	aat	tat	tct	ttt	3434
Ile -	Lys	Glu	Ile	Glu	Lys	Glu	Gln	Tyr	Ile	Ser	His	Asn	Tyr	Ser	Phe	
		1090				;	1095				1	100				
tca	att	gtt	gat	atg	ata	ttt	tta	aat	cct	tta	gat	aaa	aat	atg	ata	3482
Ser	Ile	Val	Asp	Met	Ile	Phe	Leu	Asn	Pro	Leu	Asp	Lys	Asn	Met	Ile	
1	105					1110				:	1115					
												,				
gat.	cat	gta	ata	aaa	caa	aat	aaa	cat	caa	tat	tta	att	act	tat	gaa	3530
Asp	His	Val	Ile	Lys	Gln	Asn	Lys	His	Gln	Tyr	Leu	Ile	Thr	Tyr	Glu	
1120	)				1125					1130				:	1135	
gat	aat	act	ata	ggt	ggt	ttt	tct	aca	cat	ttc	aat	aat	tat	tta	ata	3578
Asp	Asn	Thr	Ile	Gly	Gly	Phe	Ser	Thr	His	Phe	Asn	Asn	Tyr	Leu	Ile	
		•		1140	)				1145					1150		
	•															
qaa	aat	aat	tat	att	aca	aaa	cat	aac	tta	tat	gtt	cat	aat	att	tat	3626
			Tyr													
			1155			_		1160					1165			
tta	tet	aat	тас	r cca	att	gaa	cat	qca	tct	ttt	aag	gat	caa	caa	gaa	3674
			. Glu													
DCG	JC.	1170			,		1175					1180				
		11/	,			•			•							
a+-	- ۲۰		a atç	. ~~4		. +~+	ant	ctt	ato	aat	aga.	att	aaa	aat	tat	3722
			a acç s Met													
			s met	. Asį	אי י			. neu			1195	-10	~, ~		- ] -	7
	1189	•				1190	,									
									<b>h</b>		.+	++~+	2227	+		3770
ctt	aaa	a aa	t aat	t cct	t aca	a cga	itgta	ıaga	caaa	cata	La C	CCCC	aaad	-		30

Leu Lys Asn Asn Pro Thr 1200 1205

tarrettett teatacetta atgegracaa taaaatatat atceaaatat attetateeg 3830

tacgcttttt ttttttttt tttaattgtt atttttgtat at

3872

<210> 4

<211> 1205

<212> PRT

<213> Plasmodium falciparum

<400> 4

Met Ile Phe Asn Tyr Val Phe Phe Lys Asn Phe Val Pro Val Val Leu

1 5 10 15

Tyr Ile Leu Leu Ile Ile Tyr Ile Asn Leu Asn Gly Met Asn Asn Lys 20 25 30

Asn Gln Ile Lys Thr Glu Lys Ile Tyr Ile Lys Lys Leu Asn Arg Leu
35 40 45

Ser Arg Lys Asn Ser Leu Cys Ser Ser Lys Asn Lys Ile Ala Cys Leu 50 . 55 60

Phe Asp Ile Gly Asn Asp Asp Asn Arg Asn Thr Thr Tyr Gly Tyr Asn 65 70 75 80

Val Asn Val Lys Asn Asp Asp Ile Asn Ser Leu Leu Lys Asn Asn Tyr 85 90 95

Ser Asn Lys Leu Tyr Met Asp Lys Arg Lys Asn Ile Asn Asn Val Ile 100 105 110

Ser Thr Asn Lys Ile Ser Gly Ser Ile Ser Asn Ile Cys Ser Arg Asn 115 120 125

Gln Lys Glu Asn Glu Gln Lys Arg Asn Lys Gln Arg Cys Leu Thr Gln 130 135 140

Cys His Thr Tyr Asn Met Ser His Glu Gln Asp Lys Leu Ala Asn Asp

			16												
145			150				155							•	
	•									•				•	
Asn	Asn	Arg	Asn	Asn	Lys	Lys	Asn	Phe	Asn	Leu	Leu	Phe	Ile	Asn	Tyr
		•		165					170					175	

Phe Asn Leu Lys Arg Met Lys Asn Ser Leu Leu Asn Lys Asp Asn Phe 180 185 190

Phe Tyr Cys Lys Glu Lys Lys Leu Ser Phe Leu His Lys Ala Tyr Lys 195 200 205

Lys Lys Asn Cys Thr Phe Gln Asn Tyr Ser Leu Lys Arg Lys Ser Asn 210 215 220

Arg Asp Ser His Lys Leu Phe Ser Gly Glu Phe Asp Asp Tyr Thr Asn 225 230 235 240

Asn Asn Ala Leu Tyr Glu Ser Glu Lys Lys Glu Tyr Ile Thr Leu Asn 245 250 255

Asn Asn Asn Lys Asn Asn Asn Asn Lys Asn Asn Asn Asn Asn Asn Asn 260 265 270

Asp Asn Asn Asp Tyr Asn Asn Asn Ser Cys Asn Asn Leu Gly Glu 275 280 285

Arg Ser Asn His Tyr Asp Asn Tyr Gly Gly Asp Asn Asn Asn Pro Cys 290 295 300

Asn Asn Asn Asn Asp Lys Tyr Asp Ile Gly Lys Tyr Phe Lys Gln Ile 305 310 315 320

Asn Thr Phe Ile Asn Ile Asp Glu Tyr Lys Thr Ile Tyr Gly Asp Glu 325 330 335

Ile Tyr Lys Glu Ile Tyr Glu Leu Tyr Val Glu Arg Asn Ile Pro Glu 340 345 350

Tyr Tyr Glu Arg Lys Tyr Phe Ser Glu Asp Ile Lys Lys Ser Val Leu 355 360 365

Phe Asp Ile Asp Lys Tyr Asn Asp Val Glu Phe Glu Lys Ala Ile Lys

Glu Glu Phe Ile Asn Asn Gly Val Tyr Ile Asn Asn Ile Asp Asn Thr Tyr Tyr Lys Lys Glu Asn Ile Leu Ile Met Lys Lys Ile Leu His Tyr .. 410 Phe Pro Leu Leu Lys Leu Ile Asn Asn Pro Ser Asp Leu Lys Lys Leu Lys Lys Gln Tyr Leu Pro Leu Leu Ala His Glu Leu Lys Ile Phe Leu Phe Phe Ile Val Asn Ile Thr Gly Gly His Phe Ser Ser Val Leu Ser Ser Leu Glu Ile Gln Leu Leu Leu Leu Tyr Ile Phe Asn Gln Pro Tyr Asp Asn Val Ile Tyr Asp Ile Gly His Gln Ala Tyr Val His Lys Ile 485. Leu Thr Gly Arg Lys Leu Leu Phe Leu Ser Leu Arg Asn Lys Lys Gly Ile Ser Gly Phe Leu Asn Ile Phe Glu Ser Ile Tyr Asp Lys Phe Gly Ala Gly His Ser Ser Thr Ser Leu Ser Ala Ile Gln Gly Tyr Tyr Glu 

Ala Glu Trp Gln Val Lys Asn Lys Glu Lys Tyr Gly Asn Gly Asp Ile 

Glu Ile Ser Asp Asn Ala Asn Val Thr Asn Asn Glu Arg Ile Phe Gln 

Lys Gly Ile His Asn Asp Asn Asn Ile Asn Asn Ile Asn Asn Asn 590 -

Asn Tyr Ile Asn Pro Ser Asp Val Val Gly Arg Glu Asn Thr Asn Val

														•	
Pro	Asn 610	Val	Arg	Asn	Asp	Asn 615	His	Asn	Val	Asp	Lys 620	Vạl	His	Ile	.Ala
				<b>6</b> 1	C1	T ou	<b>ጥ</b> ኮ ~	Gliv	Glv	Mot	Ala	ī.en	Glu	Ala	Leu
11e 625	Ile	Gly	Asp	GIÀ	Gly 630	Leu	Inr		GIY	635	,	Dea	GIU	ALG	640
Asn	Tyr	Ile	Ser	Phe	Leu	Asn	Ser	Lys	Ile	Leu	Ile	Ile	Tyr	Asn	Ąsp
				645					650 °					655	
Asn	Gly	Gln			Leu	Pro	Thr	Asn 665	Ala	Val	Ser	Ile	Ser 670	Gly	Asn
			660			•					· •••				•
Arg	Pro	Ile 675		Ser	Ile	Ser	Asp 680	His	Leu	His	Tyr	Phe 685	Val	Ser	Asn
-1	<b>61</b>	• • •		21-	Gly	Non.	) en	Tue	Len	Ser	T.ve	Asn	Ala	Lvs	Glu
TTE	690		Asn	Ala	GIY	695	Maii	пуэ		502	700			-,-	
Asn	Asn	Ile	Phe	. Glu	Asn	Leu	Asn	Tyr	Asp	Tyr	Ile	Gly	Val	Val	Asn
705					710			· .		715			٠		720
Gly	Asn	Asr	n Thr		Glu	Leu	Phe	Lys		Leu	Asn	Asn	Ile		
				725	•				730					735	
Asn	Lÿs	Lev	ı Lys 74(		Ala	Thr	Val	Leu 745	His	Val	Arg	Thr	Lys 750		Ser
						_			<b>-</b> 1-	0	. 714	7 0.0	·	Cor	· Tle
Asn	Asp	Ph:		e Asr	n Ser	: Lys	760		ше	Ser	ile	765		Jer	110
Lve	Lv:	s As	n Gl	u Ile	e Phe	e Pro	Phe	a Asp	Thr	Thr	: Ile	Leu	ı Asn	Gly	Ası
-,-	77				*	775				•	780				
Ile	e Hi	s Ly	s Gl	u Ası	n Lys	s Ile	e Glu	ı Glu	Gļu	Lys	Asn	val	l Ser	Ser	Sei
78	5	-			790	)				795	5				800
Th:	r Ly	s Ty	r As		l Ası	n Asr	Ly:	s Asr			a Asr	ASI	n Asp		
				80					810					.81	
G1	u Il	e Il	e Ly	s Ty	r Gl	u Ası	p Me	t Phe	Ser	Lys	s Glu	1 Th	r Phe	th:	As

830 .

Ile Tyr Thr Asn Glu Met Leu Lys Tyr Leu Lys Lys Asp Arg Asn Ile 835 840 845

825

The Phe Leu Ser Pro Ala Met Leu Gly Gly Ser Gly Leu Val Lys Ile 850 855 860

Ser Glu Arg Tyr Pro Asn Asn Val Tyr Asp Val Gly Ile Ala Glu Gln 865 870 875 880

His Ser Val Thr Phe Ala Ala Ala Met Ala Met Asn Lys Lys Leu Lys 885 890 895

Ile Gln Leu Cys Ile Tyr Ser Thr Phe Leu Gln Arg Ala Tyr Asp Gln 900 905 910

Ile Ile His Asp Leu Asn Leu Gln Asn Ile Pro Leu Lys Val Ile Ile 915 920 925

Gly Arg Ser Gly Leu Val Gly Glu Asp Gly Ala Thr His Gln Gly Ile 930 935 940

Tyr Asp Leu Ser Tyr Leu Gly Thr Leu Asn Asn Ala Tyr Ile Ile Ser 945 950 955 960

Pro Ser Asn Gln Val Asp Leu Lys Arg Ala Leu Arg Phe Ala Tyr Leu 965 970 975

Asp Lys Asp His Ser Val Tyr Ile Arg Ile Pro Arg Met Asn Ile Leu 980 985 990

Ser Asp Lys Tyr Met Lys Gly Tyr Leu Asn Ile His Met Lys Asn Glu 995 1000 1005

Ser Lys Asn Ile Asp Val Asn Val Asp Ile Asn Asp Asp Val Asp Lys
1010 1015 1020

Tyr Ser Glu Glu Tyr Met Asp Asp Asp Asn Phe Ile Lys Ser Phe Ile
025 1030 1035 1040

Gly Lys Ser Arg Ile Ile Lys Met Asp Asn Glu Asn Asn Asn Thr Asn

1045

1050

1055

Glu His Tyr Ser Ser Arg Gly Asp Thr Gln Thr Lys Lys Lys Val 1060 1065 1070

Cys Ile Phe Asn Met Gly Ser Met Leu Phe Asn Val Ile Asn Ala Ile 1075 1080 1085

Lys Glu Ile Glu Lys Glu Gln Tyr Ile Ser His Asn Tyr Ser Phe Ser 1090 1095 1100

Ile Val Asp Met Ile Phe Leu Asn Pro Leu Asp Lys Asn Met Ile Asp 105 1110 1115 1120

His Val Ile Lys Gln Asn Lys His Gln Tyr Leu Ile Thr Tyr Glu Asp 1125 1130 1135

Asn Thr Ile Gly Gly Phe Ser Thr His Phe Asn Asn Tyr Leu Ile Glu 1140 1145 1150

Asn Asn Tyr Ile Thr Lys His Asn Leu Tyr Val His Asn Ile Tyr Leu 1155 1160 1165

Ser Asn Glu Pro Ile Glu His Ala Ser Phe Lys Asp Gln Gln Glu Val 1170 1175 1180

Val Lys Met Asp Lys Cys Ser Leu Val Asn Arg Ile Lys Asn Tyr Leu 185 1190 1195 1200

Lys Asn Asn Pro Thr 1205

<210> 5

<211> 3147

<212> DNA

<213> Plasmodium falciparum

<220>

<221>, CDS

<222> (199)..(2670)

<400> !										٠.			•	1	
tttcat	tttt	cttta	ccca	c at	atata	atat	ata	tata	tat	aata	tata	ta 1	tataa	tatta	60
					•									+ ~+ ~+	120
tatätt	tgat	atatg	attt	a aa	attg	taac	ata	aaaa	aaa	taat	tata		aaacc	icgigi	
atacat				- +-	++=+		'fat	tatt	att	++++	++++	tt 1	tttt	cataa	180
atacat	CLCC	aacat	ataa	a la	CLAC										
tgcctg	aata	accac	caaa	atg	agt	tat a	ata	aaa	aga	ctg	att	ctt	ttt	atg	231
- 9 9					Ser										
				1				5					10	•	-
										,	•				
tta ct															279
Leu Le	u Ph	e Tyr	Ser	His	Val	Lys	Ile	Lys	Lys	Leu	Phe	Ile	Lys	Ile	
		15					20					25		ν.	
															. 297
tct aa															327
Ser As			Ile	Phe	Phe	35	GIU	ATA	гÀг	гåз	40	GIY	цуз	цуз	
	3	0				33					10				
gaa tt	-c ++	t ctt	+++	tta	cta	aat	ata	aaa	aaa	aat	agc	caa	cag	aaa	375
Glu Ph															
	15				50					55					
aaa a	ct ta	t cat	att	acc	aaa	agg	aat	acc	ata	aat	aaa	agt	gat	ttt	423
Lys T	hr Ty	r His	Ile	Thr	Lys	Arg	Asn	Thr	Ile	Asn	Lys	Ser	Asp	Phe	
60		•		65					70					75	
	•													. 0	
tta t															471
Leu T	yr S	er Leu	ı Leü	Asn	Glu	Glu	Gly			Ser	Lys	Lys			
			80					85					90		
								226	250	2+2	C22	aat	- ata	aaa	519
		ta aaa eu Ly:													
Lys A	sn L			GIU	GIU	гÀ2	100			110	02	10!	_	-7-	
		9:	5				100								
222 +		at as	- tat	act		aaa	tat	aaa	agg	ctc	cca	ac	a cga	gaa ·	567
		ys Gl													
· Lys I		ys G1 10			,~	115		-			120				
	•												٠.		
gta c	jtt a	tt gg	a aat	: gtt	. aaa	att	gga	gga	aat	aat	aaa	at	a gct	att	615
													e Ala		

																•
caa	act	atg	gct	ągc	tgt	gat	aca	aga	aat	gta	gaa	gaa	tgt	gťa	tat	663
Gln	Thr	Met	Ala	Ser	Cys	Asp	Thr	Arg	Asn	Va1	Glu	Glu	Cys	Val	Tyr	
140					145					150					155	
		•									-					
caa	att	aga	aaa	tgt	aaa	gat	ttg	ggt	gct	gac	att	gta	agg	ttg	act	711
			Lys													•
		_	_	160	_				165					170	•	•
												. •			•	
att	caa	qqa	gtt	caa	gaa	gca	caa	gct	agt	tat	cat	att	aaa	gaa	aaa	759
			Val													÷
		-	175			•		180					185			
						•				•	٠.	•			•	•
tta	tta	tct	gaa	aat	qta	aat	atc	cca	tta	gta	gca	gat	att	cat	ttt	807
															Phe	
-;-		190					195					200			•	
•							•									
aat	cct	aaa	ata	act	tta	atg	qca	gct	gat	gtg	ttt	gaa	aaa	att	cga	855
			Ile													
	205	-4-		,		210			_		215					
				•				•					•			
ata	aat	cca	gga	aat	tat	att	gat	gga.	aga	aaa	aaa	tgg	ata	gat	aaa	903
			Gly													
220			1		225		•			230		_			235	
						•										
att	tat	äaa	act	aaa	gaa	gaa	ttt	qat	gaa	ggg	aaa	tta	ttt	ata	aaa	951
			Thr													
		-1-		240					245					250		
				•												
gaa	aaa		gta	cca	tta	att	даа	aaa	tat	aaa	aga	tta	aat	aga	gca	999
			Val													
GIU	, pys	2116	255		, Dec			260			5		265			
			233							•						
			: gga			cat		tee	ctt	tca	tet	cga	gta	tta	tca	1047
			gga Gly													
116	Arc			Thi	ASI	nis			Tiệu	Jer		280		200	002	
		270	,				275	,				. 200				
,0	. 🗆												+++	<b>a</b> 2~	+++	1095
			a gat													1033
Тух			Ası	Th:	r Pro			, met	. val	GIU			rne	GIU	FIIG	
	285	5				290	) :				295					

										э.					•	
Ct	gat	tta	tgt	att	gaa	aac	aat	ttt	tac	aat	ctt	gtt	ttt	tct	atg	1143
Ser	Asp	Leu	Cys	Ile	Glu	Asn	Asn	Phe .	Tyr	Asn	Leu	Val	Phe	Ser		
300					305					310	•	•		•	315	
			•		•											
aaa	gct	tct	aat	gct	tat	gtt	atg	ata	caa	tct	tat	aga	tta	tta	gta	1191
			Asn													
			•	320				. •	325					330		
								٠	•							
tct	aaa	caa	tat	gaa	aga	aat	atg	atg	ttc	cct	ata	cat	tta	gga	gtt	1239
Ser	Lys	Gln	Tyr	Glu	Arg	Asn	Met	Met	Phe	Pro	Ile	His	Leu	Gly	Val	
	-		335					340					345			•
٠.							•						•			
aca	gaa	gca	ада	ttt	ggt	gat	aat	gga	aga	ata	aaa	tct	tat	tta	ggt <sub>.</sub> .	1287
			Gly													
		350					355					360	•			
					•											•
ata	gga	tct	: tta	tta	tat	gat	ggt.	ata	gga	gat	acc	att	cgt	ata	tcc	1335
			Leu													
	365					370					375					
	•									•						
tta	aca	а па:	a gat	cct	t ac	gaa	gag	tta	act	cct	tgt	aaa	aaa	tta	gtt	1383
			ı Asp													
380		. 02.	. 1.DF		385					390					395	
500					-				•							
		t tt:	a aag	T 22	a aga	ata	ttt	tat	aat	gaa	aat	ttt	aaa	gaa	gat	1431
			u Lys													
			u	40				-	405					410		
		:														
221	11.	a ++	a aa	2 22	+ 221	t daa	ato	gat	acc	aaa	aat	cta	tta	aat	ttt	1479
															Phe	
ASI	. 61	u be			II , ASI	. 020		420		•			425			
			41									•				
				<b>.</b>		+ ++·	- 221	· aat	ata		a aaa	aga	aat	αta	gaa	1527
															Glu	∴.
GI	n GT			r Ai	g As	II FIII	435			,	<i>-</i>	440				
		43	30				43.					110			•	
												عنہ ۔		٠+-	Late:	1575
aa	a aa	it aa	it aa	t gt	a tt	a ca	t gaa	gaç	, tg0	act	. ata	ı ggt	. aal	. y.a	gta'	10.0
Ly	s As	n As	sn As	n Va	l Le			ı Glı	ı Cys	s In			, asi	ıval	Val	
	4 4	15			٠.	45	0			,	455	•	•			
							: •						, 			1422
															aat	1 <del>6</del> 23
Th	r T	le I.	vs G1	las La	en Gl	u As	p Se	r Lei	u Gl	n Il	e Ph	e Lys	a Asp	) Le	ı Asn'	

475

	tta	gaa	gta	gat	tca	aat	gga	aat	ttg	aaa	aag	gga <sub>.</sub>	gċc	aaa	aca	act	1671
	Leu	Glu	Val	Asp	Ser	Asn	Gly	Asn	Leu	Lys	Lys	Gly	Ala	Lys	Thr	Thr	
					480	٠.			••	485		•	•		490		
		• •				•	•										
	gat	atg	gtt	att	ata	aat	gat	ttt	cat	aat	ata	aca	aat	tta	gga	aaa '	1719
	-					Asn											.•
	_			495					500			:		505	. •		
	•		•							•	•		٠.				
	aaa	act	gtg	gat	aaa	tta	atg	caa	gtg	gga	att	aat	ata	gta	gtt	caa	1767
						Leu											••
•			510			·		515			:	•	.520				
										. •							
	tat	gaa	cca	cat	aat	ata	.gaa	ttt	ata	gaa	aaa	atg	gaa	cca	aat	aat	1815
	Tyr	Glu	Pro	His	Asn	Ile	Glu	Phe	Ile	Glu	Lys	Met.	Glu	Pro	Asn	Asn	
		525					530		•			535					
									•							-	•
	gat	aat	aat	aat	aat	aat	aat	aat	aat	aat	ata	tta	ttt	tat	gtg	gat	1863
	Asp	Asn	Asn	Asn	Asn	Asn	Asn	Asn	Asn	Asn	Ile	Leu	Phe	Tyr	Val	Asp	
	540					545					550					555	
				•													
	ata	aaa	aat	att	atg	aac	agt	tca	gaa	aaa	aat	att	aaa	tta	agt	aat	1911
	Ile	Lys	Asn	Ile	Met	Asn	Ser	Ser	Glu	Lys	Asn	Ile	Lys	Leu	Ser	Asn	
					560					565					570		
	tct	aaa	gga	tat	gga	tta	att	tta	aac	gga	aaa	gaa	gat	ata	caa	acc	1959
	Ser	Lys	Gly	Tyr	Gly	Leu	Ile	Leu	Asn	Gly	Lys	Glu	Asp	Ile	Gln	Thr	
		•		575			•		580	•				585			
	ata	aaa	aaa	ata	aaa	gaa	tta	aat	cgt	cgt	cct	tta	ttc	att	cta	tta	2007
	Ile	Lys	Lys	Ile	Lys	Glu	Leu	Asn	Arg	Arg	Pro	Leu	Phe	Ile	Leu	Leu	
			590	)				595	i				600				•
																•	
	aaa	tca	gat	aac	ata	tat	gaa	cat	gta	tta	ata	acc	aga	aga	att	aat	2055
	Lys	Ser	Asp	Asn	Ile	Tyr	Glu	His	. Val	Leu	Ile	Thr	Arg	Arg	Ile	Asn	
		605	•				610	)		•		615					
	gaa	ctt	: tta	a caa	tco	tta	aat	ata	aat	ata	cct	tat	ata	cat	tat	gtt	2103
																Val	
	620					625					630				•	635	

	****	00/1/	دب							25						
gat	att	aat	tca	aac	aat.	tat	gat	gat	ata	tta	gtt	aat	tca	aca	tta	2151
Asp	Ile	Asn	Ser	Asn	Asn	Tyr	Asp	Asp	Ile	Leu	Val	Asn :	Ser,	Thr	Leu	
•				640					645	•				650	٠.	
											• •					• :
tat	qca	gga	agt	tat	ttg	atg	gat	tta	atg	ggg	gat	ggt	ctt	att	gtt	2199
Tvr	Ala	Glv	Ser	Cvs	Leu	Met	Asp	Leu	Met	Gly	Asp	Gly ·	Leu	Ile	Val	•
		2	655					660					665			
											•		•			
226	ata	act	aat	gat	att	ctt	aca	aat	aaa	aaa	aag	ạta	gaa	aca	aaa	2247
Aen	Val	Thr	Asn	Asp	Val	Leu	Thr	Asn.	Lys	Lys	Lys	Ile	Glú	Thr	Lys	
,		670					675			-		680			•	
٠		0,0												•		
	ast	~==	222	. caa	σаа	σta	gaq	gaa	gag	gga	aac	aat	aaa	gat	att .	2295
T.	yac	Clu	Tue	. gau	Glu	Val	Glu	Glu	Glu	Gly	Asn	Asn	Lys	Asp	Ile	
Tyr			гъ	GIU	GIU	690				•	695		٠.			
_	685					0,50								, ·		
							-	tta	aat	tca	ttt	tta	aca	tta	aat	2343
cat	aga	CTT	. ככי	gago	aga	. 900	nla	Leu	Asn	Ser	Phe	Leu	Thr	Leu	Asn	•
		Lei	Let	u Ser			, ALC	neu	11011	710			•		715	
700					705	•				, 10						
												ast	tat	ata	acc	2391
att	tta	caa	a ga	t aca	a aga	ata	cgt	tta	Dba		mh-	gat	Tur	Tle	Ala	
Ile	Let	ı Glı	n As	p Th	r Arç	, Ile	e Arq	Leu			Int	Asp	ıyı	730		
				72	0				725	•				,50		
															222	2439
tgo	cca	a tc	t tg	t gg	a aga	a act	t tta	a ttt	: aat	ata	caa	gaa	mb-	, acc	aaa	2.07
Ċys	Pr	o \$e	r Cy	s Gl	y Ar	g Th	r Le			, 1Te	e GTI	GIU			Lys	
			73	5				740	)				745	•		•
																2407
aaa	a at	t at	g aa	a tt	a ac	a gg	g ca	c tta	a aaa	a ggo	gtt	aaa	att	gca	gtc	2487
Lys	s Il	e Me	t Ly	s Le	u Th	r Gl	y Hi	s Le	ı Ly	s Gl	y Val	Lys	Ile	e Ala	Val.	
		75	0				75	5				760	)		,	
										•						
ate	g gg	a to	t at	tt gt	t aa	t gg	t at	a gg	a ga	a at	g gca	a gat	gca	a cat	ttt	2535
Me	t Gl	v Cv	s I	le Va	ıl As	n Gl	y Il	e Gl	y Gl	u Me	t Ala	a Asp	Ala	a His	s Phe	
	76					77					77					
	•	-							٠,				,			
~~	+ +-		· 	at e	יד מר	a co	t aa	a aa	a at	t ga	t tt	a tat	: ta	t ggi	t aaa	2583
99	. n.	"	y	yı aç	,	a Di	o f.	's Lv	s Il	e As	p Le	u Tyr	Ty	r Gl	Lys	
		YI V	al G	τλ <b>2</b> 6			)	,		79					795	
78	·				78	, ,					-					•
							:	·• ~-		~-	a . ac	t tai	t aa	t aa	a tto	2631
ga	g t	ta g	ta g	aa a	ga aa	at a	ta co	ec ga	ug ga	ia ya		- ca	. ya	n tu	a ttg s Leu	
G1	lu L	eu V	al G	lu A	rg A	sn I	le P	ro Gl	.u G1	.u Gl	u Al	a cy:	s no	P PA	s Leu	

800

810

ata gaa tta att aaa aaa cat aac aaa tgg aaa gat cca taaattgaat 2680
Ile Glu Leu Ile Lys Lys His Asn Lys Trp Lys Asp Pro
815 820

<210> 6

<211> 824

<212> PRT ·.

<213> Plasmodium falciparum

<400> 6

Met Ser Tyr Ile Lys Arg Leu Ile Leu Phe Met Leu Leu Phe Tyr Ser 1 5 10 15

His Val Lys Ile Lys Lys Leu Phe Ile Lys Ile Ser Asn Val Asn Ile 20 25 30

Phe Phe Ala Glu Ala Lys Lys Asn Gly Lys Lys Glu Phe Phe Leu Phe 35 40 45

Leu Leu Asn Ile Lys Lys Asn Ser Gln Gln Lys Lys Thr Tyr His Ile 50 55 60

Thr Lys Arg Asn Thr Ile Asn Lys Ser Asp Phe Leu Tyr Ser Leu Leu

.

75

80

Asn Glu Glu Gly Asn Ser Ser Lys Lys Glu Tyr Lys Asn Leu Lys Asp
85 90 95

70

Glu Glu Lys Tyr Asn Ile Ile Gln Asn Ile Lys Lys Tyr Cys Glu Cys 100 105 110

Thr Lys Lys Tyr Lys Arg Leu Pro Thr Arg Glu Val Val Ile Gly Asn 115 120 125

Val Lys Ile Gly Gly Asn Asn Lys Ile Ala Ile Gln Thr Met Ala Ser 130 135 140

Cys Asp Thr Arg Asn Val Glu Glu Cys Val Tyr Gln Ile Arg Lys Cys
145 150 155 160

Lys Asp Leu Gly Ala Asp Ile Val Arg Leu Thr Val Gln Gly Val Gln 165 170 175

Glu Ala Gln Ala Ser Tyr His Ile Lys Glu Lys Leu Leu Ser Glu Asn 180 185 190

Val Asn Ile Pro Leu Val Ala Asp Ile His Phe Asn Pro Lys Ile Ala 195 200 205

Leu Met Ala Ala Asp Val Phe Glu Lys Ile Arg Vál Asn Pro Gly Asn 210 215 220

Tyr Val Asp Gly Arg Lys Lys Trp Ile Asp Lys Val Tyr Lys Thr Lys 225 230 235 240

Glu Glu Phe Asp Glu Gly Lys Leu Phe Ile Lys Glu Lys Phe Val Pro 245 250 255

Leu Ile Glu Lys Cys Lys Arg Leu Asn Arg Ala Ile Arg Ile Gly Thr 260 265 270

Asn His Gly Ser Leu Ser Ser Arg Val Leu Ser Tyr Tyr Gly Asp Thr 275 280 285

Pro Leu Gly Met Val Glu Ser Ala Phe Glu Phe Ser Asp Leu Cys Ile

Glu Asn Asn Phe Tyr Asn Leu Val Phe Ser Met Lys Ala Ser Asn Ala Tyr Val Met Ile Gln Ser Tyr Arg Leu Leu Val Ser Lys Gln Tyr Glu Arg Asn Met Met Phe Pro Ile His Leu Gly Val Thr Glu Ala Gly Phe Gly Asp Asn Gly Arg Ile Lys Ser Tyr Leu Gly Ile Gly Ser Leu Leu Tyr Asp Gly Ile Gly Asp Thr Ile Arg Ile Ser Leu Thr Glu Asp Pro Trp Glu Glu Leu Thr Pro Cys Lys Lys Leu Val Glu Asn Leu Lys Lys Arg Ile Phe Tyr Asn Glu Asn Phe Lys Glu Asp Asn Glu Leu Lys Asn Asn Glu Met Asp Thr Lys Asn Leu Leu Asn Phe Glu Glu Asn Tyr Arg Asn Phe Asn Asn Ile Lys Lys Arg Asn Val Glu Lys Asn Asn Asn Val Leu His Glu Glu Cys Thr Ile Gly Asn Val Val Thr Ile Lys Glu Leu Glu Asp Ser Leu Gln Ile Phe Lys Asp Leu Asn Leu Glu Val Asp Ser Asn Gly Asn Leu Lys Lys Gly Ala Lys Thr Thr Asp Met Val Ile Ile Asn Asp Phe His Asn Ile Thr Asn Leu Gly Lys Lys Thr Val Asp Lys

Leu Met Gln Val Gly Ile Asn Ile Val Val Gln Tyr Glu Pro His Asn

515	520	525
313	520	

Ile Glu Phe Ile Glu Lys Met Glu Pro Asn Asn Asn Asn Asn Asn 530 540

Asn Asn Asn Asn Ile Leu Phe Tyr Val Asp Ile Lys Asn Ile Met 545 550 555 560

Asn Ser Ser Glu Lys Asn Ile Lys Leu Ser Asn Ser Lys Gly Tyr Gly 565 570 575

Leu Ile Leu Asn Gly Lys Glu Asp Ile Gln Thr Ile Lys Lys Ile Lys
580 585 590

Glu Leu Asn Arg Arg Pro Leu Phe Ile Leu Leu Lys Ser Asp Asn Ile
595 600 605

Tyr Glu His Val Leu Ile Thr Arg Arg Ile Asn Glu Leu Leu Gln Ser 610 615 620

Leu Asn Ile Asn Ile Pro Tyr Ile His Tyr Val Asp Ile Asn Ser Asn 625 630 635 640

Asn Tyr Asp Asp Ile Leu Val Asn Ser Thr Leu Tyr Ala Gly Ser Cys 645 650 655

Leu Met Asp Leu Met Gly Asp Gly Leu Ile Val Asn Val Thr Asn Asp 660 665 670

Val Leu Thr Asn Lys Lys Lys Ile Glu Thr Lys Tyr Asp Glu Lys Glu 675 680 685

Glu Val Glu Glu Glu Gly Asn Asn Lys Asp Ile His Arg Leu Leu Ser 690 695 700

Arg Val Ala Leu Asn Ser Phe Leu Thr Leu Asn Ile Leu Gln Asp Thr 705 710 715 720

Arg Ile Arg Leu Phe Lys Thr Asp Tyr Ile Ala Cys Pro Ser Cys Gly
725 730 735

Arg Thr Leu Phe Asn Ile Gln Glu Thr Thr Lys Lys Ile Met Lys Leu

PCT WELTORGANISATION FÜR GEISTIGES EIGENTUM
Internationale Bürd
INTERNATIONALE ANMELDUNG VERÖFFENTLICHT NACH DEM VERTRAG ÜBER DIE INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES PATENTWESENS (PCT)

(51) Internationale Patentklassifikation 7:

C12N 9/90, 9/10, 9/12, C12Q 1/48

(11) Internationale Veröffentlichungsnummer:

WO 00/17233

**A3** 

(43) Internationales Veröffentlichungsdatum:

30, März 2000 (30.03.00)

(21) Internationales Aktenzeichen:

PCT/EP99/07055

(22) Internationales Anmeldedatum:

22. September 1999

(22.09.99)

(30) Prioritätsdaten:

198 43 279.8 199 23 567.8 22. September 1998 (22.09.98) DE

DR 21. Mai 1999 (21.05.99)

(71)(72) Anmelder und Erfinder: JOMAA, Hassan [DE/DE]; Breslauer Strasse 24, D-35398 Gießen (DE).

(74) Anwälte: PANTEN, Kirsten usw.; Reichel und Reichel, Parkstrasse 13. D-60322 Frankfurt am Main (DE).

(81) Bestimmungsstaaten: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DK, DM, EE, ES, FL, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO Patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), eurasisches Patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), europäisches Patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI Patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Veröffentlicht

Mit internationalem Recherchenbericht.

(88) Veröffentlichungsdatum des internationalen Recherchenberichts: 25. Mai 2000 (25.05.00)

(54) Title: GENES OF THE 1-DESOXY-D-XYLULOSE BIOSYNTHETIC PATHWAY

(54) Bezeichnung: GENE DES 1-DESOXY-D-XYLULOSE-BIOSYNTHESEWEGS

The invention relates to the 1-desoxy- D-xylulose- 5-phosphate reductoisomerase gene, the 1-desoxy- D-xylulose- 5-phosphatesynthase gene and the gcpE gene of the 1-desoxy- D-xylulose biosynthetic pathway and to their use for transforming vectors, host organisms and plants and for determining substances that inhibit this biosynthetic pathway.

(57) Zusammenfassung

Die vorliegende Erfindung betrifft das 1-Desoxy- D-xylulose- 5-phosphatreduktoisomerase -Gen, das 1-Desoxy- D-xylulose-5-phosphat- Synthase- Gen und das gcpE-Gen des 1-Desoxy- D-xylulose- Biosynthesewegs und ihre Verwendung zur Transformation von Vektoren, Wirtsorganismen und Pflanzen und zur Bestimmung von Stoffen, die diesen Biosyntheseweg inhibieren.

#### LEDIGLICH ZUR INFORMATION

Codes zur Identifizierung von PCT-Vertragsstaaten auf den Kopfbögen der Schriften, die internationale Anmeldungen gemäss dem PCT veröffentlichen.

LS	Lesotho	SI	Slowenien
LT	Litauen	SK	Slowakei
LU	Luxemburg	SN	Senegal
LV	Lentland	SZ	Swasiland
MC	Monaco	TD	Tschad
MD	Republik Moldau	TG	Togo
MG	Madagaskar	TJ	Tadschikistan
MK	Die ehemalige jugoslawische	TM	Turkmenistan
WERE	Republik Mazedonica	TR	Türkei
241	•		Trinidad und Tobago
			Ukraine
			Uganda
			Vereinigte Staaten von
		w	Amerika
			Usbekistan
			Vietnam
	***************************************		
NO	Norwegea		Jugoslawien
NZ	Neusceland	zw	Zimbabwe
PL	Polen		
PT	Portugal	•	
RO	Rumanien		
RU	Russische Föderation		
SD	Sudan		
SE	Schweden		
	Singapur		
	PL PT RO RU SD	ML Mali MN Mongolei MR Mauretanien MW Malawi MX Mexiko NE Niger NL Niederlande NO Norwegen NZ Neuseeland PL Polen PT Portugal RO Rumanien RU Russische Föderation SD Sudan SE Schweden	ML Mali TT MN Mongolei UA MR Mauretanien UG MW Malawi US MX Mexiko NE Niger UZ NL Niederlande VN NO Norwegen YU NZ Neusceland ZW PL Polen PT Portugal RO Rumlanien RU Russische Föderation SD Sudan SE Schweden

PCT/EP 99/07055

	CATION OF BUBLECT MATTER C12N9/90 C12N9/10 C12N9/12	C1201/48	
According to (	international Patent Classification (IPC) or to both national classifica	ation and IPC	
B. FIELDS 8			
	sumeritation searched (classification system followed by classification C12N C12Q	on symbols)	
		the state of the state of the state of	orbed .
	on searched other than minimum documentation to the extent that a	•	
Bectronic da	da base consulted during the International search (name of data be	se and, where practical, search terms used	'
1			
1			
C. DOCUME	ENTS CONSIDERED TO BE RELEVANT		
Category *	Chation of document, with indication, where appropriate, of the re-	levers passages	Relevant to claim No.
E.L	WO 99 52938 A (HASSAN JOMAA)	·	1,2,4,
	21 October 1999 (1999-10-21)		8-12, 16-18
	see SeqID's;Priority of inv. s W09952938 and PCTEP99/07055 and DE19816196.4,DE19825585.3,DE1982 9831637.2,DE19831639.9 or DE1983 be invalid; A.4C(4) PC.	present in 8097.1,DE1	.0 .0
	_	-/	
1	·	· Y	,
		•	. 7
			·
1		•	
1 .			
1.			
			10
	erther documents are fisted in the continuation of box C.	Peters family members are liste	
1	categories of cited documents:	"I later document published after the im or priority date and not in conflict with	
con	ment defining the general state of the last which is not aldered to be of particular relevance	cited to understand the principle or a invention	about the state of
	or document but published on or after the International g date	"X" document of particular relevance; the cannot be considered novel or cannot be considered novel or cannot be considered novel.	
"L" docus	ment which may throw doubts on priority claim(s) or on is cited to establish the publication date of another	involve an inventive step when the	cislmed invention
cita	tion or other special reason (as specified) iment referring to an oral disclosure, use, exhibition or	carmot be considered to involve as	none other such doout-
oth	er means	ments, such combination being down in the sat.	OUR DE Persons and
	rment published prior to the international filing date but " or than the priority date claimed	"&" document member of the same peter	
Date of t	he actual completion of the international sourch	Date of mailing of the international e	earch report
	7 March 2000	23/03/2000	
Name ar	nd mailing address of the ISA	Authorized officer	
	European Patent Office, P.B. 5818 Patentiaan 2 NL - 2290 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl. Fax: (+31-70) 340-3016	Hoekstra, S	

Irik Sonel Application No PCT/EP 99/07055

(Continue	ntion) DOCUMENTS CONSIDERED TO BE RELEVANT	Relevent to claim No.
ategory *	Citation of document, with indication, where appropriate, of the relevant passages	Michael in result see
χ .	PUTRA ET AL: "Incorporation of	15
	'2,3-'13!C2!- and '2,4-'13!C2!-D-1-Deoxyxylulose into	
	ubiquinone of Escherichia coli via the Mevalonate-Independent pathway for	
•	Isoprenoid Biosynthesis" TETRAHEDRON LETTERS,NL,ELSEVIER SCIENCE	
*	PUBLISHERS, AMSTERDAM, vol. 39, no. 39, 1998, pages 23-26-26,	
	XP002116676 ISSN: 0040-4039	
٠.	figure 1	
x .	KUZUYAMA ET AL: "Direct formation of	15
	2-C-Methyl-D-Erythritol 4-phosphate from 1-Deoxy-D-Xylulose 5-phosphate	
	Reductoisomerase, a new enzyme in the non-mevalonate pathway to isopentenyl	
	diphosphate" TETRAHEDRON LETTERS, NL, ELSEVIER SCIENCE PUBLISHERS, AMSTERDAM,	
	vol. 39, no. 39, 1998, pages 4509-4512-44512, XP002116675	
	ISSN: 0040-4039 figure 1	
P,X	SCHWENDER. J. ET AL.: "Cloning and	1,9-12
•	heterologous expression of a cDNA encoding 1-deoxy-d-xylulose-5-phosphate reductoisomerase of Arabidopsis thaliana"	
•	FEBS LETTERS, vol. 455, July 1999 (1999-07), pages 140-144, XP002132424	
	the whole document	4
P,A	DE 197 52 700 A (HOECHST SCHERING AGREVO GMBH) 2 June 1999 (1999-06-02) the whole document	1-12
A	LANGE ET AL: "A family of transketolases that directs isoprenoid biosynthesis via a	1–18
	mevalonate-independent pathway* FASEB JOURNAL, US, FED. OF AMERICAN SOC. FOR EXPERIMENTAL BIOLOGY, BETHESDA, MD,	
	vol. 95, March 1998 (1998-03), pages 2100-2104, XP002116672 ISSN: 0892-6638	
	the whole document	
P;X	EMINY DATABASE: "AC: EF111813" PLASMODIUM FALCIPARUM 1-DEOXY-D-XYLULOSE 5-PHOSPHATE REDUCTOISOMERASE,	1,9-12
	11 January 1999 (1999-01-11), XP002132425 see : Scores	
1		

tota Jonel Application No PCT/EP 99/07055

		PC1/EP 99/0/055	
C.(Continue Cetegory *	tion) DOCUMENTS CONSIDERED TO BE RELEVANT  Citation of document, with indication, where appropriate, of the relevant pessages	Relevent to	o dalm No.
P,X	TREMBL DATABASE: "AC: 096693" PLASMODIUM FALCIPARUM 1-DEOXY-D-XYLULOSE 5-PHOSPHATE REDUCTOISOMERASE, 1 May 1999 (1999-05-01), XP002132426 see: Scores	1,	9–12
X.	SPRENGER ET AL: "Identification of a thiamin-dependent synthase in Escherichia coli required for the formation of the 1-deoxy-D-xylulose 5-phosphate precursor to isoprenoids, thiamin, and pyridoxol" FASEB JOURNAL, US, FED. OF AMERICAN SOC. FOR EXPERIMENTAL BIOLOGY, BETHESDA, MD, vol. 94, November 1997 (1997-11), pages 12857-12862, XP002116674 ISSN: 0892-6638 figure 2	2,	9-12
A	figure 1	1	
P,X	TREMBL DATABASE: "AC: QZ8HO" CHLAMYDIA PNEUMONIAE GCPE PROTEIN,	3	,9–12
	1 May 1999 (1999-05-01), XP002132427		
	see : scores		•
	·		•
		*	
		, i	
		· · · ·	
	·	·	
		·	
1	·	. [	
1			
ļ	4		
	·	·	
	•		

Form PCT/ISA/210 (continuation of second sheet) (Ady 1992

information on patent family members

trik Jonel Application No PCT/EP 99/07055

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9952938 A	21-10-1999	DE 19825585 A DE 19828097 A DE 19831637 A AU 4120899 A AU 4481699 A WO 9952515 A WO 9966875 A WO 0004031 A WO 0003699 A	21-10-1999 30-12-1999 27-01-2000 01-11-1999 01-11-1999 21-10-1999 29-12-1999 27-01-2000 27-01-2000
DE 19752700 A	02-06-1999	DE 29800547 U JP 11169186 A	08-04-19 <b>99</b> 29-06-19 <b>99</b>

PCT/EP 99/07055

	•	101/21 33/	
a klassifi IPK 7	C12N9/90 C12N9/10 C12N9/12	C12Q1/48	
Nach der Inte	rmetionelen Petertidassifikation (IPK) oder nach der nætionalen Klassifi	kation and der IPK	
	CHIEFTE GEBIETE		
	or Mindestprütstoff (Klassifikationssystem und Klassifikationssymbole)		
IPK 7	C12N C12Q		
Recherchiert	ze aber nicht zum Mindestprüfstoff gehörende Veröffertlichungen, sowe	it diese unter die recherchierten Gebiete t	olen
Während de	internationalen Recherche konauftlene elektroniache Dalenbank (Nam	ie der Datenbank und evil. venwendete S	uchbegriffe).
	•		
C ALS W/F	SENTLICH ANGESEHENE UNTERLAGEN	<u> </u>	
Kategorie*	Bezeichnung der Veröffentlichung, soweit erforderlich unter Angebe d	ier in Betracht kommenden Teile	Betr. Anspruch Nr.
E,L	WO 99 52938 A (HASSAN JOMAA) 21. Oktober 1999 (1999–10–21)		1,2,4, 8-12, 16-18
	Siehe SeqID's; Priority of inv. sha W09952938 and PCTEP99/07055 and pr DE19816196.4, DE19825585.3, DE198280 9831637.2, DE19831639.9 or DE198316 be invalid; A.4C(4) PC.	esent in 197.1,DE1	
	-/	<b>/</b>	
[ ] w	ettere Veröffentlichungen sind der Fortsetzung von Feld C zu	Siehe Anheng Peteritamille	· ·
Beconde "A" Verdy aber "E" ättere Ann "L" Verdy and soll ause "O" Verd eine "P" Verd den	thermen  The Kategorien von angegebenen Veröffentlichungen  The Kategorien von angegebenen Veröffentlichungen  The Kategorien von angegebenen Veröffentlichungen  The Colument, das jedoch erst am oder nach dem Internationalen  teldedatum veröffentlicht worden ist  tertlichung, die geelgnet ist, einen Prioritätssanspruch zweifehaft er-  tern zu lassen, oder durch die das Veröffentlichungedatum einer  eren im Recherchenbericht genannten Veröffentlichung belegt werden  oder die aus einem anderen besonderen Grund angegeben ist (wie  geführt)  flentlichung, die sich auf eine mündliche Offenbarung.	T* Spätere Veröffertlichung, die nach den oder dem Prioritätsdatum veröffertlich Ammeldung richt kollidert, sondem nie Erfindung zugrundellegenden Prinzipi Theorie angegeben ist "X" Veröffertlichung von besonderer Bede kann eilein aufgrund dieser Veröffertlich in Zufatet betruhend betreitlichen ist füt gleit beruhend betreitlichen ist füt gleitlichen ist gl	ar zum Verständrie des der  soder der ihr zugrundeliegenden  utung; die beanspruchte Erlindung  ichtung nicht als neu oder auf  achtet werden  utung; die beanspruchte Erlindung  keit beruhend betrachtet  it einer oder mehreren anderen  n Verbindung gebracht wird und  n nahell egend lat  n Patentfamilie let
	7. März 2000	23/03/2000	
Name ur	nd Postanechrift der Internationalen Recherchenbehörde Europäisches Patentarrit, P.B. 5818 Patentiaan 2 NL – 2280 HV Rijewijk Tel. (+31–70) 340–2040, Tx. 31 651 epo ni,	Bevolimächögter Bedlensteller Hoekstra, S	
1	Ear (491-70) 940-9016	1	

Formblett PCT/ISA/210 (Bleft 2) (Auf 1992)

PCT/EP 99/07055

(Fortsetn	ING) ALS WESENTLICH ANGESEHENE UNTERLAGEN		
ategorie*	Bezeichnung der Veröffentlichung, sowett erforderlich unter Angebe der in Betracht komme	nden Telle	Betr. Anspruch Nr.
	PUTRA ET AL: "Incorporation of		15
	'2.3-'13!C2!- and		
1	'2,4-'13!C2!-D-1-Deoxyxylulose into ubiquinone of Escherichia coli via the	•	
	Mevalonate-Independent pathway for		
	Isoprenoid Biosynthesis"		
	TETRAHEDRON LETTERS, NL, ELSEVIER SCIENCE		
	PUBLISHERS, AMSTERDAM,	•	
	Bd. 39, Nr. 39, 1998, Seiten 23-26-26, XP002116676		
	ISSN: 0040-4039	•	· ·
	Abbildung 1		
	KUZUYAMA ET AL: "Direct formation of		15
(	KUZUYAMA ET AL: "Direct formation of 2-C-Methyl-D-Erythritol 4-phosphate from		
	1-Deoxy-D-Xylulose 5-phosphate	•	
	Reductoisomerase, a new enzyme in the		
	non-mevalonate pathway to isopentenyl	•	]
	diphosphate" TETRAHEDRON LETTERS, NL, ELSEVIER SCIENCE		
	PUBLISHERS. AMSTERDAM,	•	
	Bd. 39, Nr. 39, 1998, Seiten		
	4509-4512-44512, XP002116675		
	ISSN: 0040-4039 Abbildung 1		
			1 0-12
P,X	SCHWENDER, J. ET AL.: "Cloning and		1,9-12
	heterologous expression of a cDNA encoding 1-deoxy-d-xylulose-5-phosphate		
	reductoisomerase of Arabidopsis thaliana*		
	FEBS LETTERS.		
	Bd. 455, Juli 1999 (1999-07), Seiten		
	140-144, XP002132424 das ganze Dokument		
P,A	DE 197 52 700 A (HOECHST SCHERING AGREVO	· .	1-12
	GMBH) 2. Juni 1999 (1999-06-02)		
	das ganze Dokument		
۸	LANGE ET AL: "A family of transketolases		1-18
l	that directs isoprenoid biosynthesis via a		
	mevalonate-independent pathway		
	FASEB JOURNAL, US, FED. OF AMERICAN SOC. FOR EXPERIMENTAL BIOLOGY, BETHESDA, MD,		
1	Bd. 95, März 1998 (1998-03), Seiten		
	2100-2104, XP0021166 <b>72</b>		1 .
	ISSN: 0892-6638		
	das ganze Dokument		
P.X	EMINY DATABASE: "AC: EF111813"		1,9-12
	PLASMODIUM FALCIPARUM 1-DEOXY-D-XYLULUSE		
	5-PHOSPHATE REDUCTOISOMERASE, 11. Januar 1999 (1999-01-11), XP002132425		
	Siehe: Scores		
1	- Orente, oddres		
1	-/		
Ţ	1		

PCT/EP 99/07055

		701/21 33/0/033
(Fortsetz (etegorie*	mg) ALS WESENTLICH ANGEBEHENE UNTERLAGEN  Bezeichnung der Veröffentlichung, soweit erforderlich unter Angebe der in Betracht komm	menden Telle Betr. Anspruch Nr.
Ρ,Χ	TREMBL DATABASE: "AC: 096693" PLASMODIUM FALCIPARUM 1-DEOXY-D-XYLULOSE 5-PHOSPHATE REDUCTOISOMERASE, 1. Mai 1999 (1999-05-01), XP002132426 Siehe: Scores	1,9-12
K A	SPRENGER ET AL: "Identification of a thiamin-dependent synthase in Escherichia coli required for the formation of the 1-deoxy-D-xylulose 5-phosphate precursor to isoprenoids, thiamin, and pyridoxol" FASEB JOURNAL, US, FED. OF AMERICAN SOC. FOR EXPERIMENTAL BIOLOGY, BETHESDA, MD, Bd. 94, November 1997 (1997-11), Seiten 12857-12862, XP002116674 ISSN: 0892-6638 Abbildung 2 Abbildung 1	2,9–12
A P,X	TREMBL DATABASE: "AC: QZ8HO" CHLAMYDIA PNEUMONIAE GCPE PROTEIN, 1. Mai 1999 (1999-05-01), XP002132427 Siehe: scores	3,9–12
		·

Angeben zu Veröffentlichungen, die zur seiben Patentiemlie gehören

Int. Jonelee Aktenzeichen
PCT/EP 99/07055

Im Recherchenbericht; angeführtes Patentdokument	Datum <b>der</b> Veröffentlichung	. Mitglied(er) <b>der</b> Patentiamilie	Datum der Veröffentlichung
WO 9952938 A	21-10-1999	DE 19825585 A DE 19828097 A DE 19831637 A AU 4120899 A AU 4481699 A WO 9952515 A WO 9966875 A WO 0004031 A WO 0003699 A	21-10-1999 30-12-1999 27-01-2000 01-11-1999 01-11-1999 21-10-1999 29-12-1999 27-01-2000 27-01-2000
DE 19752700 A	02-06-1999	DE 29800547 U JP 11169186 A	08-04-19 <b>99</b> 29-06-19 <b>99</b>